Incentives and Efficiency of Pension Systems*

Ali Shourideh Maxim Troshkin[†]

December, 2017

We study efficiency and incentive costs of social insurance and redistribution when retirement is endogenous. We characterize Pareto optima, show the forces determining optimal retirement ages, and derive the properties of optimal retirement distortions. It is optimal for a pension system to reward later retirement independent of whether efficient retirement ages increase or decrease in productivity. When individual heterogeneity and the parameters of status-quo policies are calibrated to U.S. income taxes, Social Security, individual earnings, hours, and retirement ages, optimal pensions generate not only significant welfare gains but also aggregate output gains.

JEL codes: E62, H21, H55. *Keywords*: optimal fiscal policy, social insurance, redistribution, endogenous retirement.

1 Introduction

Pension systems are by far the largest component of social insurance as well as a means of redistribution in most countries. An optimal system should provide productive workers with incentives to fully realize their potential while providing benefits to the individuals experiencing low productivity. On the one hand, standard production efficiency implies that more-productive workers should supply more labor – later retirement of the more-productive could be an aspect of that. On the other hand, one way more-productive workers are incentivized is with leisure – sufficiently high needs for incentives would require earlier retirement of more-productive workers.

We study the trade-off between efficiency and incentives in order to evaluate the constrained efficiency of the U.S. Social Security system and to assess the welfare consequences of optimal dynamic nonlinear policies, including the optimal dependence of

*We thank V.V. Chari, Stephen Coate, Helmuth Cremer, Mikhail Golosov, Philippe Choné, Roozbeh Hosseini, Mark Huggett, Larry Jones, Nadia Karamcheva, Dirk Krueger, Luigi Pistaferri, Tom Sargent, Aleh Tsyvinski, Şevin Yeltekin and numerous seminar and conference participants for valuable input and Tirupam Goel for excellent research assistance. This research was in part supported by the Institute for the Social Sciences.

[†]Shourideh: ashourid@andrew.cmu.edu, Carnegie Mellon University, Tepper School of Business. Troshkin: troshkin@cornell.edu, Cornell University, Department of Economics.

pensions on retirement age. This requires a way to qualitatively characterize as well as to quantify constrained efficiency in a model of individual work and retirement choices, with salient features of status quo taxes and benefits and with plausible individual elasticities at both the hours and retirement margins.[1]

We start from a life-cycle setup with agents who are ex ante heterogeneous in productivity, which evolves with age, and in fixed costs of working, which are allowed to be correlated with productivity. Fixed costs make individual budget sets non-convex and, combined with hump-shaped productivity-age profiles, make it individually optimal to choose to retire at some age, while choosing a hump-shaped labor supply profile during the working life. The agents are privately informed about their productivities and their fixed costs. Re-writing the planner's mechanism design problem in terms of virtual productivities and virtual fixed costs (accounting for information rents), we prove a novel equivalence result giving a straightforward characterization of constrained-efficient allocations with endogenous retirement. That allows us to show the forces determining optimal retirement ages and to derive the properties of a class of income tax and pension policies that implement constrained efficiency, but are also close to the U.S. status quo in a way that can be made precise. In that sense the optimal policies we characterize require relatively small changes to the status quo.

We then turn to a positive version of the setup, where individuals observed in the U.S. data take the status-quo income taxes and the Social Security as given and make individually optimal choices, which potentially are not constrained-efficient. We calibrate the setup using combined micro-level data from the Health and Retirement Survey (HRS) and the U.S. Panel Study of Income Dynamics (PSID). We show that the positive setup can replicate key features of reality both internally (in terms of targeted moments such as retirement ages and labor hours) and externally (in terms of the income distribution and the elasticity at the retirement margin). A principle contribution is this ability to reproduce salient features of reality and then, fixing the estimated parameters, to quantify constrained-efficient labor supply, retirement ages, and optimal policies in a unified setup. We then extend this analysis to overlapping generations, to accounting for intergenerational transfers, and to heterogeneous life spans.

We document three main findings. First, we show that rewarding later retirement is optimal without regard for whether optimal retirement ages increase or decrease in productivity. This is because it is always costlier to provide incentives by distorting retirement ages rather than by distorting hours worked. We show that optimal retirement

[1] By retirement we will mean a decision to stop supplying labor. A decision to claim retirement benefits, such as Social Security, will be separate and treated distinctly from the labor force exit decision.

distortions are low, in particular lower than labor distortions, and it is optimal for a pension to reward later retirement in order to counteract the incentives for earlier retirement created by any positive income taxes. In other words, it is optimal for pensions to directly depend on the retirement age so that the present value of lifetime benefits rises with the age of retirement, independent of whether or not more-productive agents should retire older.

Second, we show that the optimal retirement ages of higher-lifetime-earnings individuals are significantly higher than in the U.S. data, and also much higher than those of their less-productive peers. Our equivalence result allows a type of sufficient statistic expression for constrained-efficient retirement ages to increase or decrease in productivity. That relationship is driven by individual differences in productivities relative to fixed costs, augmented by the Frisch elasticity of labor supply. It allows us to evaluate how far from constrained-efficient the observed retirement ages are. For example, workers in the top half of the lifetime earnings distribution should optimally retire on average at the age of 69.8 in our baseline, while in the U.S. data their average retirement age is 66.9. On the other hand, workers in the bottom half should optimally retire at 61.7, while in the data they do so at 67.6.

Third, we quantify how much stronger the incentives for later retirement would need to be in an optimal system vs. the existing U.S. Social Security and tax systems. This is achieved by increases in pension benefits in response to delaying retirement. Optimal marginal (with respect to retirement age) pension benefits are positive, quantitatively significant, and increase with earnings: from 4.5 percent per year of delayed retirement at the bottom of the lifetime earnings distribution to 18 percent at the top.[2] We show that optimal pensions with such incentives create large aggregate welfare gains but also aggregate output gains of up to 1.7 percent, unlike in much of the optimal taxation literature. We also show that reforms that do not account for individualized incentives (e.g., requiring retirement at the time of pension benefit claiming, or uniformly increasing retirement age) keep the economy away from constrained-efficiency.

Relation to previous literature. Constrained efficiency with extensive margins of labor supply, from various perspectives, is at the center of a related literature with a theoretical focus (e.g., Saez (2002)), recently focusing on the optimality of history-dependent distortions when life-cycle considerations are introduced (e.g., Cremer, Lozachmeur, and Pestieau (2004), Michau (2014), Choné and Laroque (2015)). The contribution of our equivalence result is that – unlike in the previous studies – it enables a straightforward,

[2]Marginal pension benefits are well defined in a continuous time setting. Delayed retirement credit in the U.S. Social Security system can be thought of as their analogue in the U.S. data.

complete characterization explaining the main forces in terms of standard intuitive trade-offs. This facilitates a surprising and new finding that optimal pension system implementations provide incentives for later retirement independent of whether efficient retirement ages increase or decrease in productivity.

Our findings also contribute to a growing empirically-motivated literature bringing theoretical constrained-efficient wedges to estimable distributions and elasticities (e.g., Saez (2001), Golosov, Troshkin, and Tsyvinski (2011), Saez and Stantcheva (2016)). We generalize commonly found tax formulas by connecting standard labor wedges to retirement wedges through estimable elasticities and distributions. An advantage of this connection is a type of sufficient statistic for the efficient retirement age, a novel finding that is shown to be quite useful.

On the quantitative side, most recent studies of optimal redistributive policies largely find that increasing policy distortions (vs. the status quo) significantly improves welfare but generally sacrifices aggregate output, e.g., Weinzierl (2011), Farhi and Werning (2013). In contrast, we show that output may not need to be sacrificed if efficient incentives for retirement are taken into account. Our estimation of the environment parameters uses a mixed identification strategy following recent literature on idiosyncratic consumption and labor choices, e.g., Low, Meghir, and Pistaferri (2010); our use of estimated fixed effects from earnings regressions as types and parts of our exposition follow Low and Pistaferri (2015). We focus on permanent shocks following a literature that finds that most of the welfare gains come not from insurance against temporary shocks but from the provision of social insurance against permanent shocks (e.g., Huggett and Parra (2010)). A related literature finds that permanent shocks similar to the ones we focus on account for most of the variation in lifetime earnings and lifetime utility (e.g., Huggett, Ventura, and Yaron (2011)). Our framework with fixed costs follows recent work applying non-convex budget sets as a source of retirement decision in life-cycle settings, e.g., Rogerson and Wallenius (2013).

An important complementary approach to these general questions is to study policy reforms within a set of parametrically restricted policy instruments as in, e.g., Conesa, Kitao, and Krueger (2009) in the context of dynamic taxation and Huggett and Parra (2010) in the context of Social Security. More recently Golosov et al. (2013) restrict the parametric set to stylized versions of status-quo Social Security and fix taxes, resulting in quite different optima and welfare effects from what we find. Our analysis contributes to that line of research by informing which properties are salient in the choice of the parametric sets of policies.

The rest of the paper is organized as follows. Section 2 describes the general environment, from both the positive and the normative perspectives. Qualitative properties of the constrained-efficient allocations and optimal policy are characterized in Section 3. Section 4 discusses the construction of a quantitative environment and Section 5 quantifies constrained optima and optimal policies. Section 6 concludes.

2 The Environment

2.1 Normative setup

Consider a continuum of individuals born at $t = 0$ who live a continuous interval of time until $t = \bar{T}$. Each individual is born with a type, $\theta \in \Theta \equiv \left[\underline{\theta}, \bar{\theta}\right]$, drawn at $t = 0$ from a distribution $F(\theta)$ with $F'(\theta) = f(\theta) > 0$ for all θ. The type affects the idiosyncratic productivity-age profile $\varphi(t, \theta)$: an individual of type θ who chooses at age t to work $l(t, \theta)$ hours produces $y(t, \theta) = \varphi(t, \theta)\, l(t, \theta)$ units of output. Assume φ is twice continuously differentiable and inverse U-shaped, i.e., for each $\theta \in \Theta$ there exists an age t^* such that $\frac{\partial \varphi(t,\theta)}{\partial t} > 0$ for all $t < t^*$ and $\frac{\partial \varphi(t,\theta)}{\partial t} < 0$ for all $t > t^*$.

The type also affects the idiosyncratic fixed utility cost of working $\eta(t, \theta)$: an individual who chooses $l(t, \theta) > 0$ pays fixed utility cost $\eta(t, \theta)$ in addition to standard continuous disutility from work. Assume η is continuously differentiable and non-decreasing with age.

The preferences of an individual of type θ are given by

$$\int_0^{\bar{T}} e^{-\rho t} \left[u\left(c\left(t, \theta\right)\right) - v\left(\frac{y\left(t, \theta\right)}{\varphi\left(t, \theta\right)}\right) - \eta\left(t, \theta\right) \mathbf{1}\left\{y\left(t, \theta\right) > 0\right\} \right] dt, \tag{1}$$

where ρ is a subjective discount factor, $c(t, \theta)$ denotes consumption at age t, u is strictly concave, increasing, and satisfies Inada conditions, v is strictly convex with $v'(0) = 0$, and $\mathbf{1}\{\cdot\}$ is an indicator function. The presence of η makes the total disutility of working non-convex, implying non-convex individual budget sets. A non-convex budget set can lead an individual to optimally choose a discontinuous drop in hours at some age, even with continuous hours choice and a preference for smoothing leisure over life.[3]

An allocation for a cohort of individuals, $\left(c\left(t, \theta\right), y\left(t, \theta\right)\right)_{\theta \in \Theta, t \in [0, \bar{T}]}$, is feasible if

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_0^{\bar{T}} e^{-rt} c\left(t, \theta\right) dt dF\left(\theta\right) + H_0 \leq \int_{\underline{\theta}}^{\bar{\theta}} \int_0^{\bar{T}} e^{-rt} y\left(t, \theta\right) dt dF\left(\theta\right) + rK_0, \tag{2}$$

[3]See, e.g., Rogerson and Wallenius (2013). They also review empirical evidence that retirement appears as abrupt transitions from full-time work to not working in the U.S. micro data. In the online Appendix C we show similar behavior in a pooled sample of the HRS and the PSID individuals.

where r is the interest rate, $H_0 \equiv \int_0^{\bar{T}} e^{-rt} H_t dt$ is the present value of the government revenue requirement (net outflow of income from the cohort), K_0 is initial capital, with both K_0 and H_0 given.

The individuals are privately informed about their productivities and fixed costs. Thus any social insurance cannot contract directly on them but remains otherwise unrestricted, e.g., to be arbitrarily nonlinear or age dependent. To study the problem of a government seeking optimal social insurance, we will characterize the mechanism design problem arising from this information asymmetry. Following standard arguments, the revelation principle guarantees the sufficiency of considering direct mechanisms: individuals report their types to a fictitious planner who chooses allocations subject to incentive compatibility, i.e., for all $\theta, \hat{\theta} \in \Theta$

$$\int_0^{\bar{T}} e^{-\rho t} \left[u\left(c\left(t,\theta\right)\right) - v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right) - \eta\left(t,\theta\right) \mathbf{1}\left\{y\left(t,\theta\right) > 0\right\} \right] dt \geq \\ \int_0^{\bar{T}} e^{-\rho t} \left[u\left(c\left(t,\hat{\theta}\right)\right) - v\left(\frac{y\left(t,\hat{\theta}\right)}{\varphi\left(t,\theta\right)}\right) - \eta\left(t,\theta\right) \mathbf{1}\left\{y\left(t,\hat{\theta}\right) > 0\right\} \right] dt, \quad (3)$$

where θ is an individual's type and $\hat{\theta}$ is the individual's report about the type. The planner's objective is to maximize a social welfare function

$$\int_{\underline{\theta}}^{\bar{\theta}} U\left(\theta\right) dG\left(\theta\right), \quad (4)$$

where $U\left(\theta\right)$ is the lifetime utility of type θ given by (1) and $G\left(\theta\right)$ is a cumulative density function, differentiable over $\left(\underline{\theta}, \bar{\theta}\right]$ with $G\left(\underline{\theta}\right) = 0$, $G\left(\bar{\theta}\right) = 1$, and $G'\left(\theta\right) = g\left(\theta\right) \geq 0$. A given exogenous motive to redistribute from higher-earning individuals to lower-earning ones is captured by $G\left(\theta\right) \geq F\left(\theta\right)$ for all $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$.[4]

An allocation is *constrained efficient* if it is a solution to the direct mechanism design problem of maximizing social welfare (4) subject to incentive compatibility (3) and feasibility (2).

Discussion and extensions. This setting is geared toward analyzing implications, particularly welfare implications, of the changes in labor supply at both hours and retirement margins in response to the efficient provision of social insurance and redistribution. In related environments the welfare gains from labor and capital policies optimized within

[4]The case of $G(\theta) = 1$ for all $\theta > \underline{\theta}$ corresponds to the Rawlsian criterion; $G\left(\theta\right) = F\left(\theta\right)$ corresponds to the Utilitarian objective. We restrict the differentiability of G to the semi-open interval to include the extremes of redistributive motives like the Rawlsian criterion.

given sets of functional forms are extensively studied (e.g., Altig et al. (2001), Conesa, Kitao, and Krueger (2009)). Those gains intuitively come from providing better incentives to save – hence increasing savings – and the resulting effects on the interest rate. Significantly less understanding, however, exists of the labor supply responses, particularly the interaction between the hours and retirement responses. To focus on these mechanisms, we follow recent literature (e.g., Best and Kleven (2013), Farhi and Werning (2013)) in representing the production technology above by an AK-type production function, which is additively separable between labor and capital and therefore allows to abstract from the savings effects.[5]

While our main discussion will focus on a single cohort, an alternative interpretation is the steady state of an overlapping generations economy. We develop an overlapping generations version of this environment in the online Appendix B where each generation is exactly as described above but with additional notation to identify the generations.

Another useful extension will be to reinterpret this setting as the problem of a planner attached to one specific generation, i.e., maximizing the welfare of the generation taking as given net intergenerational transfers. We will explore that interpretation in Section 5 where we will think of H_0 as the combined present value of government spending and net intergenerational transfers, as well as allowing heterogeneous life span.

2.2 Theoretical wedges vs. policies

Because of the incentive compatibility constraints, efficient allocations could clearly drive wedges into the individual optimality conditions. Before characterizing efficient allocations, it is useful to define the wedges and contrast them with the actual policy tools that can provide implementations. Even though the analysis will extend to the general setup, the following assumption will provide an intuitive benchmark throughout:

Assumption 1 *For all $\theta \in \Theta$: (i) $\frac{\partial}{\partial t}\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)} \geq 0$, and (ii) $\eta(t,\theta) = 0$ whenever $t < t^*$ and $\eta(t,\theta) = \eta(\theta)$ otherwise.*

Condition (i) states that relative productivity differences between types do not diminish with age.[6] This is closely related to the observations that the right tail of the income

[5]Similarly, parts of that literature also focus on permanent shocks following findings that such shocks account for most of the variation in lifetime earnings and lifetime utility (e.g., Huggett, Ventura, and Yaron (2011)) and that most of the welfare gains come not from redistribution with temporary shocks but from the provision of social insurance against permanent shocks (e.g., Huggett and Parra (2010)).

[6]In other words, productivity-age profiles weakly "fan-out". Studies estimating heterogeneous productivity profiles over the life-cycle generally find similar patterns (e.g., Altig et al. (2001), Nishiyama and Smetters (2007)).

distribution thickens with age and that more-productive individuals tend to have steeper growth earlier in life and slower decline later in life. This is compatible with most of the population becoming disabled by high enough age.

Condition (ii) is a simple way to assume that participation costs do not diminish with age.[7] This ensures that everyone joins the labor force at $t = 0$, focusing on labor force exit and related policies rather than on the issues related to entering the labor force. Within a positive version of the setup that we discuss below one approach to identifying the fixed costs empirically is from observed individual retirement ages. An identifying assumption in that case is that fixed costs are stable around the observed retirement ages – condition (ii) will allow us to first develop a simple intuition. We will later relax this assumption and calibrate the rate of change of fixed costs with age to match available estimates of elasticity at the retirement margin.

At every age, leisure for each type is affected by individual choices of whether to work and, if so, how much to work; for each type θ there will exist a retirement age – denote it $R(\theta)$ – such that type θ chooses $y\left(t,\theta\right) > 0$ for $t < R\left(\theta\right)$ and $y\left(t,\theta\right) = 0$ for $t \geq R\left(\theta\right)$.[8] We show it formally in the online Appendix A and write allocations as $\left(c\left(t,\theta\right)_{t\in[0,\bar{T}]}, y\left(t,\theta\right)_{t\in[0,R(\theta)]}, R\left(\theta\right)\right)_{\theta\in\Theta}$. One can then define *labor wedge*, $\tau_y\left(t,\theta\right)$, and *retirement wedge*, $\tau_R\left(\theta\right)$, by the following optimality conditions:

$$\left(1-\tau_y\left(t,\theta\right)\right)u'\left(c\left(t,\theta\right)\right) = v'\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)\frac{1}{\varphi\left(t,\theta\right)}, \tag{5}$$

$$\left(1-\tau_R\left(\theta\right)\right)y\left(R\left(\theta\right),\theta\right)u'\left(c\left(R\left(\theta\right),\theta\right)\right) = v\left(\frac{y\left(R\left(\theta\right),\theta\right)}{\varphi\left(R\left(\theta\right),\theta\right)}\right) + \eta\left(R\left(\theta\right),\theta\right). \tag{6}$$

When both wedges are zero, equations (5) and (6) are simply laissez faire individual optimality conditions. That is, an efficient allocation that drives a non-zero wedge between the marginal rate of substitution and the marginal rate of transformation distorts the individually optimal consumption-labor choice in (5). Similarly, a wedge in (6) between the marginal utility of income and the marginal disutility of output, which includes the fixed cost, distorts the individual decision about the retirement age $R\left(\theta\right)$. The two distortions reflect efficient incentives constrained by the information asymmetry in individual productivities and fixed costs.

[7] It is sufficient to assume $\partial\eta\left(\theta,t\right)/\partial t \geq 0$ throughout. This is easy to see, for example, in the proof of the existence of a retirement age in the online Appendix A. Intuitively, this captures, for example, the deterioration of health with age, making the individuals not only less productive but also increasing their fixed costs of participating in the labor force.

[8] See, e.g., Cremer, Lozachmeur, and Pestieau (2004), Choné and Laroque (2015).

It is not difficult to see that the wedges of an efficient allocation can have infinitely many actual government policy tools implementing them – a standard property of the dynamic mechanism design approach to optimal policy. We focus on a class of policies that not only include implementations of constrained efficiency but also include stylized status quo. Next we define this class in the context of the positive setup, then prove in the next section that the class also contains implementations of efficiency and explicitly characterize their qualitative properties. We later show that with the status-quo policies from this class the positive setup describes reality quantitatively well, allowing the quantitative analysis of efficient policies within a coherent framework.

2.3 Positive setup

Consider a class of government policies consisting of an individual income tax function $\mathcal{T}$ and the present value of net retirement benefits b. Taking these policies as given, a type-θ individual maximizes life-time utility (1) subject to the present value budget constraint

$$\int_0^{\bar{T}} e^{-rt} c\left(t,\theta\right) dt = \int_0^{R(\theta)} e^{-rt}\left[y\left(t,\theta\right) - \mathcal{T}\left(t, y\left(t,\theta\right)\right)\right] dt + b\left(R\left(\theta\right), Y\left(\theta\right)\right), \tag{7}$$

where $Y\left(\theta\right) \equiv Y\left(y\left(t,\theta\right)_{t\in[0,R(\theta)]}\right)$ is a measure of lifetime earnings. Note that $\mathcal{T}$ is potentially age dependant, but history independent as it is a function of only the current realization of income, while b is a function of the history of incomes. Note also a distinction between the age of claiming benefits and the age of retirement: while per-period benefits depend on both, given an actuarially fair policy the present value of benefits b is a function of the age of retirement.[9]

Individual optimality conditions can still be written as equations (5) and (6) but with the wedges replaced by the interactions in the policy tools as follows:

$$\tau_y\left(t,\theta\right) = \mathcal{T}_y\left(t, y\left(t,\theta\right)\right) - e^{rt}\delta_{y(t,\theta)} Y\left(\theta\right) b_Y\left(R\left(\theta\right), Y\left(\theta\right)\right),$$

[9]That is, we allow each individual to have access to risk-free savings and borrowing so that the instantaneous budget constraint for a given θ is

$$c\left(t\right) + \dot{a}\left(t\right) = \mathbf{1}_{\{t\leq R\}}\left(y\left(t\right) - \mathcal{T}\left(t, y\left(t\right)\right)\right) + \mathbf{1}_{\{t>S\}} b\left(R,Y\right)\left(e^{-rS} - e^{-r\bar{T}}\right)/r + ra\left(t\right),$$

where a is the level of individual asset holdings and then $b\left(R,Y\right)\left(e^{-rS} - e^{-r\bar{T}}\right)/r$ is the per-period benefit claimed starting at age S. Such asset holdings capture, for instance, employer-provided pensions or other tax-deferred accounts.

$$\tau_R(\theta)\, y(R(\theta),\theta) = \mathcal{T}(R(\theta), y(R(\theta),\theta)) - e^{rR(\theta)} b_R(R(\theta), Y(\theta)) \\ - e^{rR(\theta)} \left(\int_{R(\theta)}^{\bar{T}} e^{-rt} dt \right) \delta_R Y(\theta)\, b_Y(R(\theta), Y(\theta)),$$

where $\delta_{y(t)} Y$ is the Fréchet derivative of Y with respect to $y(t)$ and $\delta_R Y$ is with respect to R. Evidently there is a direct connection between the specific properties of the tools in this class of policies and our definitions of the wedges. It gives an intuitive meaning to the wedges. The labor wedge here becomes the balance between the distortions from the marginal income tax and from the marginal benefits with respect to lifetime earnings. The retirement wedge trades off the total tax burden at retirement against the change in benefits coming from adjusting retirement age.

3 Qualitative Properties of Efficient Allocations

To first isolate the forces that are fundamental for the results, we abstract in this section from risk aversion and discounting and set $H_0 = K_0 = 0$. For concreteness, assume quasi-linear utility with $v(l) = \psi l^{1+1/\varepsilon} / (1+1/\varepsilon)$, where ψ is a strictly positive constant and $\varepsilon \in (0,\infty)$ is a Frisch (intensive) elasticity of labor supply. Since the allocation of consumption for an individual is then indeterminate across time, assume without loss of generality that it is constant over the life-cycle, $c(\theta)$. We will relegate formal proofs to the online Appendix A and once we develop the intuition here we will relax these restrictions as well as provide in the online Appendix B formal proofs for the general setup.

3.1 Equivalence result

We start with an equivalence result that will prove useful: the mechanism design problem associated with efficient allocations can be characterized by instead considering a simpler "full-information" problem if productivities and fixed costs are re-defined to account for the information rents.

Incentive constraints (3) are equivalently a set of lifetime-utility maximization problems, one for each θ, with the choice variable $\hat{\theta}$, a report. Following the first-order approach the incentive compatibility can be written using the envelope theorem as (see, e.g., Ábrahám, Koehne, and Pavoni (2011), or Kapička (2013)):

$$U'(\theta) = \int_0^{R(\theta)} \psi \frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)} \frac{y(t,\theta)^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}} dt - \eta'(\theta)\, R(\theta) + (\eta(\theta) t^*(\theta))', \tag{8}$$

for all $\theta \in \Theta$, where $U(\theta)$ is the lifetime utility of type θ given by (1). Then the planner's problem with private information – maximizing welfare (4) subject to feasibility (2) and incentive compatibility (8) with U given by (1) – can be characterized by solving instead the problem omitting incentive compatibility (8) if productivities and fixed costs are appropriately modified:[10]

Proposition 1 *An allocation is constrained efficient if and only if it is efficient with productivity $\tilde{\varphi}$ and fixed costs $\tilde{\eta}$ given by*

$$\tilde{\varphi}(t,\theta) = \varphi(t,\theta)\left[1+\left(1+\frac{1}{\varepsilon}\right)\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)}\right]^{-\frac{\varepsilon}{1+\varepsilon}} \tag{9}$$

$$\tilde{\eta}(\theta) = \eta(\theta)\left[1-\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\eta'(\theta)}{\eta(\theta)}\right] \tag{10}$$

What gives rise to this equivalence intuitively is the fact that individuals possess private information about their types and hence a constrained-efficient allocation must allow them to collect rents on that information. Those rents effectively modify productivities and fixed costs to reflect how they are perceived by the individuals. In particular, larger relative differences in productivities, $\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)}$, require stronger incentives for more productive types and hence constrained-efficient allocations must deliver larger information rents to those individuals. The modified productivities in (9) capture exactly that, augmented by the Frisch elasticity. Analogous incentives and information-rent effect are produced by larger relative differences in fixed costs, $\frac{\eta'(\theta)}{\eta(\theta)}$. The modification in (10) accounts for that without the need to account for the hours elasticity. At the same time, stronger preferences for redistribution toward a particular type, $\frac{G(\theta)-F(\theta)}{f(\theta)}$, naturally produce the same effects for both the productivities and fixed costs.

3.2 Characterizing efficient retirement ages

One immediate benefit of the equivalence result is that it allows one to understand constrained-efficient retirement ages from standard public information trade-offs. Consider a marginal increase in retirement age $R(\theta)$. It has a mechanical effect of increasing output by $\tilde{\varphi}(R,\theta)\, l(R,\theta)$. It also has welfare effects of increasing the disutility of working by

[10] We show in the online Appendix B that with risk aversion the required modification is analogous. To reflect the redistributive motives in G it is no longer enough to compare simply to the distribution of types F since it is no longer the case that social welfare function is the only source of curvature. Hence the comparison there is with a Utilitarian motive accounting for both sources of curvature. One consequence, for instance, is that the modification is no longer degenerate in a Utilitarian case. The modification required by Proposition 1 is also related to the concept of virtual types of Myerson (1981).

$\psi l\left(R,\theta\right)^{1+1/\varepsilon}/\left(1+1/\varepsilon\right)$ and by virtual fixed cost $\tilde{\eta}\left(\theta\right)$. At the efficient $R\left(\theta\right)$ these effects must balance: output net of variable cost of hours must be equal to the virtual fixed cost. On the other hand, optimality conditions also imply that when hours are chosen efficiently output net of variable cost of hours is proportional to $\tilde{\varphi}\left(t,\theta\right)^{1+\varepsilon}$. Thus $\tilde{\varphi}\left(t,\theta\right)^{1+\varepsilon}/\tilde{\eta}\left(\theta\right)$ must be equated across types at their efficient retirement ages, i.e.,

$$\tilde{\varphi}\left(R\left(\theta\right),\theta\right)^{1+\varepsilon}/\tilde{\eta}\left(\theta\right)=\kappa$$

for some constant κ for all $\theta\in\Theta$. Differentiating with respect to θ, the implicit function theorem yields

$$R'\left(\theta\right)=-\frac{\left.\frac{\partial}{\partial\theta}\tilde{\varphi}\left(t,\theta\right)^{1+\varepsilon}/\tilde{\eta}\left(\theta\right)\right|_{t=R(\theta)}}{\left.\frac{\partial}{\partial t}\tilde{\varphi}\left(t,\theta\right)^{1+\varepsilon}/\tilde{\eta}\left(\theta\right)\right|_{t=R(\theta)}}.$$

If, for example, individual productivities are declining around retirement, it is efficient to have retirement age increase in θ if and only if $\tilde{\varphi}\left(R\left(\theta\right),\theta\right)^{1+\varepsilon}/\tilde{\eta}\left(\theta\right)$ increases in θ, and decrease otherwise:

Proposition 2 *: The constrained-efficient retirement age, $R\left(\theta\right)$, increases (decreases) in θ if and only if $\frac{\partial}{\partial\theta}\frac{\tilde{\varphi}(R(\theta),\theta)^{1+\varepsilon}}{\tilde{\eta}(\theta)}\geq 0\ \ (\leq 0)$.*

The insight here is that the efficient retirement behavior is driven by how virtual productivities $\tilde{\varphi}\left(t,\theta\right)$, augmented by the Frisch elasticity, differ across individuals relative to how different the virtual fixed costs $\tilde{\eta}\left(\theta\right)$ are. This sufficient-statistic-type expression allows one to collapse complex optimality conditions down to a function of few objects that are straightforward to interpret.[11] While not necessary for the result, Assumption 1 is useful here in focusing on the fundamental forces. It implies that the productivities are increasing in θ at a given age. Since the fixed costs can also be increasing in θ, the resulting retirement behavior must be determined by the relative change. The extent of information asymmetry makes the effective relative change more or less pronounced by adjusting information rents.

3.3 Why retirement incentives are costlier than hours incentives

Intuitively, if the retirement wedge is low relative to the labor wedge, the distortion from the income tax provides enough incentives for earlier retirement so that incentives for

[11] We show in the online Appendix B that in the general setup with risk aversion and discounting the expression is exactly the same. It also does not rely on condition (ii) in Assumption 1 as $\partial\eta\left(\theta,t\right)/\partial t\geq 0$ is sufficient. For an overview of the related general approach of sufficient statistics see, e.g., Chetty (2009).

later retirement must be provided by the pension benefits. If, on the other hand, the retirement wedge is higher than the labor wedge, the distortions from the taxes are not enough to provide efficient incentives for early-enough retirement and it must be also rewarded by the pensions.

We show next that under relevant conditions the retirement wedge is smaller than the labor wedge at retirement and that this finding is independent of whether the efficient retirement age pattern is increasing or decreasing.

Proposition 3 *The wedges implied by the constrained-efficient allocation satisfy*

$$\tau_R(\theta) = \tau_y(R(\theta), \theta) \Big/ \left(1 + \frac{1}{\varepsilon}\right) - \frac{G(\theta) - F(\theta)}{f(\theta)} \frac{\eta'(\theta)}{y(R(\theta), \theta)}. \tag{11}$$

In particular, $\tau_R(\theta) < \tau_y(R(\theta), \theta)$ *whenever* $\eta'(\theta) \geq 0$.

This finding is novel and the independence from the retirement age pattern could appear surprising or even counter-intuitive at first. To see the intuition, imagine first a simple example of $\eta'(\theta) = 0$. Compare the incentive effects of an increase in output via an adjustment in hours vs. an adjustment in the retirement age. A marginal increase in $y(R(\theta), \theta)$ by ϵ lowers the utility of working for type θ by an amount proportional to $\epsilon y(R(\theta), \theta)^{1/\varepsilon}$. The same increase in output can be achieved by increasing retirement age $R(\theta)$ by $\frac{\epsilon}{y(R(\theta),\theta)}$. This lowers the utility of retiring by an amount proportional to $\frac{\epsilon}{y(R(\theta),\theta)} \times \frac{y(R(\theta),\theta)^{1+1/\varepsilon}}{1+1/\varepsilon} = \frac{\epsilon y(R(\theta),\theta)^{1/\varepsilon}}{1+1/\varepsilon}$, which is less than from the adjustment in hours. In other words, when individuals do not differ in fixed costs, the distortions to hours are more useful in providing incentives and consequently the labor wedge at retirement is larger than the retirement wedge.

When $\eta'(\theta) \geq 0$, as will be the case with all of our estimated fixed costs, this mechanism becomes even more pronounced since increasing retirement ages naturally provide additional incentives for the more-productive types not to under-report their type.

This is also connected to standard optimal taxation formulas (e.g., the static formulas in Saez (2001) and the dynamic formulas in Golosov, Troshkin, and Tsyvinski (2011)). To show that explicitly, we extend a standard labor wedge formula accounting for the life-cycle with endogenous retirement: for all t and $\theta \in \Theta$,

$$\frac{\tau_y(t, \theta)}{1 - \tau_y(t, \theta)} = \left(1 + \frac{1}{\varepsilon}\right) \frac{G(\theta) - F(\theta)}{f(\theta)} \frac{\varphi_\theta(t, \theta)}{\varphi(t, \theta)}. \tag{12}$$

The difference here is that the labor wedge is scaled up by the relative change in the productivities, $\varphi_\theta(t, \theta) / \varphi(t, \theta)$, reflecting the additional incentives coming from the information rents discussed above. Since $\varphi_\theta(t,\theta)/\varphi(t,\theta)$ is increasing in t after the peak productivity t^*, these incentives provide an additional force beyond standard age-dependence results, increasing the labor wedge with age and making it less costly to keep the retirement distortions lower. In other words, ignoring retirement incentives would lead to optimal policy conclusions that are qualitatively different. Next we show that the main insight here carries through to implementations in the form of optimal pensions rewarding later retirement.

3.4 Pensions rewarding later retirement are optimal

To make policy implementation more intuitive, we explicitly construct the tax and pension benefit functions before deriving their qualitative properties. Given a constrained-efficient allocation $\left(c(\theta), y(t,\theta)_{t\in[0,R(\theta)]}, R(\theta)\right)_{\theta\in\Theta}$, extend $y(t,\theta)$ for $t > R(\theta)$ by defining it to be the values implied by the virtual productivities (9) if the planner were to ignore the virtual fixed costs (10). This gives a complete profile of income for all ages and individuals. Then the tax function $\mathcal{T}(t,y)$ is defined by

$$\theta = \arg\max_{\hat{\theta}} y\left(t,\hat{\theta}\right) - \mathcal{T}\left(t, y\left(t,\hat{\theta}\right)\right) - \frac{\psi}{1+1/\varepsilon}\frac{y\left(t,\hat{\theta}\right)^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}}. \tag{13}$$

Incentive compatibility of $(y(t,\theta) - \mathcal{T}(t, y(t,\theta)), y(t,\theta), \cdot)$ pins down the slope of $\mathcal{T}(t,\cdot)$ with respect to y. Then over a feasible interval the function $\mathcal{T}(t,y)$ is uniquely determined up to a constant:

Lemma 1 *Given a constrained-efficient allocation with $y(t,\theta)$ continuous and increasing in θ, there exists $\mathcal{T}(t,y)$ satisfying (13) that is unique up to a constant on $\left[y(t,\underline{\theta}), y\left(t,\bar{\theta}\right)\right]$.*

Construct now the benefits $b(R)$ by first defining

$$\hat{b}(\theta) = c(\theta) - \int_0^{R(\theta)} \left[y(t,\theta) - \mathcal{T}(t, y(t,\theta))\right] dt. \tag{14}$$

Whenever $R(\theta)$ is a one-to-one function of θ, there exists $b(R)$ such that $b(R(\theta)) = \hat{b}(\theta)$. If $R \neq R(\theta)$ for some θ, set $b(R)$ to a sufficiently small number that type θ would never choose.

Lemma 2 *Given a constrained-efficient allocation with $y(t,\theta)$ continuous and increasing in θ and $R(\theta)$ one-to-one, the policies in (13) and (14) implement the allocation, i.e., the allocation is a local optimum.*

Intuitively, the tax function $\mathcal{T}$ is constructed so that conditional on working in period t an individual of type θ earns exactly the constrained-efficient $y(t,\theta)$. Then the construction of the benefits with formula (14) means that a choice of retirement age $R = R\left(\hat{\theta}\right)$, for some $\hat{\theta}$, coincides with the choice to report $\hat{\theta}$. Given incentive compatibility constraints, however, the individual of type θ will choose $R(\theta)$. Risk neutrality makes it particularly straightforward to see this intuition because with quasi-linear utility the choice of a retirement age does not affect labor supply at each age. In the general setup with risk aversion a change in the retirement age can affect per-period consumption and hence can change the decision about hours worked at certain ages – it may become individually optimal to double deviate. Nevertheless, we show in the online Appendix B that it is sufficient to condition the benefits function on lifetime earnings to prevent such double deviations.

Properties of Optimal Pension Benefits. Assuming differentiable $b(R)$ (e.g., guaranteed by the differentiability of the allocations) individual optimality conditions imply[12]

$$y(R) - \mathcal{T}(R, y(R)) + b'(R) = \frac{\psi}{1+1/\varepsilon} \frac{y(R)^{1+1/\varepsilon}}{\varphi(R)^{1+1/\varepsilon}} + \eta(\theta),$$

implying that the retirement wedge is given by $\frac{\mathcal{T}(R,y(R))}{y(R)} - \frac{b'(R)}{y(R)}$ and hence from Proposition 3

$$\frac{\mathcal{T}(R, y(R))}{y(R)} - \frac{b'(R)}{y(R)} < \mathcal{T}_y(R, y(R)).$$

That is, pension benefits must reward later retirement, $b'(R) > 0$, whenever marginal tax rates are lower than average tax rates. Even though only the slope $\mathcal{T}_y(t,y)$ is uniquely determined at every age, for an arbitrary intercept, the marginal tax is likely to be lower than the average tax at higher incomes. Moreover, since the implementation works for any intercept, $\mathcal{T}(\cdot,\cdot)$ can always be modified so that $b'(R) > 0$.

A key novel qualitative insight here, just as in the analysis of incentives above, is that pension benefits reward later retirement independently of whether the efficient retirement ages are increasing or decreasing with productivity:

Proposition 4 *Given $\mathcal{T}$ and b implementing constrained efficiency:*

(i) Whenever average tax is at least as large as marginal tax, i.e., $b'(R(\theta)) > 0$ for all $\theta \in \Theta$ with $\mathcal{T}(R(\theta), y(R(\theta),\theta)) / y(R(\theta),\theta) \geq \mathcal{T}_y(R(\theta), y(R(\theta),\theta))$, pension benefits reward later retirement.

(ii) There always exist optimal $\hat{\mathcal{T}}$ and $\hat{b}$ implementing constrained efficiency with pension benefits rewarding later retirement, i.e., with $\hat{b}'(R(\theta)) > 0$ for all $\theta \in \Theta$.

[12] To simplify notation, we suppress here explicit dependence on θ whenever it does not jeopardize clarity.

Properties of Optimal Tax Functions. The explicit construction of $\mathcal{T}$ extends to the environment with endogenous retirement the observation that the efficient marginal taxes are generally age dependent. Formula (12) reveals, however, that age dependence is driven also by the properties of the productivity-age profiles. Specifically, whether the efficient marginal tax increases, decreases, or stays unchanged with age will be driven by how relative productivity differences evolve with age, i.e., the sign of $\frac{\partial}{\partial t}\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)}$. In other words, the rewarding of delayed retirement by optimal pensions is independent of the properties of optimal income tax functions. In particular, that finding is not affected by weather the optimal income taxes are age-dependent or not.[13]

4 Constructing Quantitative Environment

We now return to the positive version of the setup to estimate parameters of the general environment with risk aversion and discounting using U.S. microeconomic data. We assume each individual in the data takes as given the status-quo income taxes and the U.S. Social Security and maximizes lifetime utility (1) subject to the present-value budget constraint (7) as described in Section 2. The observed individually-optimal allocations are hence potentially constrained *in*efficient.

We use a mixed identification strategy following recent literature on idiosyncratic consumption and labor choices.[14] First, some parameters are fixed following existing findings or are taken directly from the observed data. The robustness is checked by changing the values of these parameters within the ranges from the literature and by using alternative definitions in the data. Second, some parameters are estimated outside the environment with ancillary statistical models. For some of these this is done to reduce the computational burden, e.g., for the productivity-age profiles. For others, estimation within the structure of the environment is not needed, e.g., for the status-quo policy functions. In the third step the remaining parameters are calibrated to match data moments, e.g., the unobservable fixed costs.

[13]This can also be seen intuitively more broadly by considering an example with productivity profiles proportional to each other, i.e., $\frac{\partial}{\partial t}\varphi(t,\theta) = 0$. Then the labor wedge is independent of age and hence so is the marginal tax function (because the marginal rate of substitution between hours worked at different ages is independent of θ and hence all individuals evaluate income at different ages the same way). As a result, variations in the marginal tax across ages cannot be useful in providing incentives. On the other hand, some degree of fanning out in productivity-age profiles can potentially make increasing marginal tax useful for some types after their peak productivity ages $t^*(\theta)$.

[14]See, e.g., Low, Meghir, and Pistaferri (2010).

Table 1: Summary statistics for the HRS, the PSID, and the pooled samples.

	HRS sample	PSID sample	Pooled sample
Individuals	971	1,116	2,087
Observations	5,788	30,751	36,539
Years of education	12.7 (3.1)	13.2 (2.5)	13.0 (2.8)
Fraction Caucasian	0.92	0.94	0.90
Fraction married	0.87	0.84	0.88
Average annual hours	2272 (635)	2129 (768)	2195 (713)
Average wage	35.0 (274.5)	23.9 (15.8)	29.1 (187.6)
Avg. retirement age, baseline definition	67.4 (5.3)	-	67.4 (5.3)
Avg. retirement age, alternative definition	68.9 (4.8)	64.8 (6.7)	66.7 (6.3)
Social Security claiming age	63.5 (2.5)	-	63.5 (2.5)

Notes: 1940-cohort males; standard deviations in parentheses.

Sources: RAND HRS, PSID data set from Heathcote, Perri, and Violante (2010).

4.1 Data

Main sources of individual data are the RAND version of the HRS, which is a cleaned and streamlined version of raw HRS files, and the PSID data set from Heathcote, Perri, and Violante (2010), who aim to carefully address a number of well-known issues in the raw data. To take advantage of both the more extensive longitudinal component and the larger retirement age sample, we construct a pooled sample of the HRS and the PSID. The number of individuals from each data set is roughly equal, with the PSID naturally providing significantly more observation per person as summarized in Table 1.[15] Summary statistics indicate close sample averages for standard demographic characteristics and hours worked as well as expectedly higher average wage for a more mature HRS sample. Our baseline is the males of the 1940 cohort. One advantage of that cohort is that it is the longest observed cohort in the HRS.[16]

The bottom three rows of Table 1 also summarize key ages for the setup. We let $t = 0$ in the model correspond to age 20 and set the baseline life span $\bar{T} = 81.6$ following Bell and Miller (2005), extending to heterogeneous life spans $T(\theta)$ later in the next section. The individual ages at which Social Security benefits are claimed, $S(\theta)$, are taken from the HRS. Retirement ages $R(\theta)$ are calculated using two definitions, a baseline and an

[15]The online Appendix C provides further details of the sample construction, the distribution of retirement ages, and how they vary in the data with the definition of retirement, with education, and by sector.

[16]Following Guvenen (2009) and Heathcote, Perri, and Violante (2010), we experimented with using all birth years and removing cohort effects with little change to the results below.

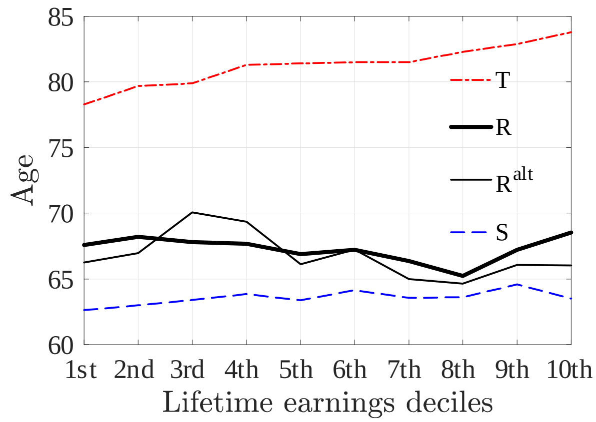

Figure 1: The U.S. microeconomic data (pooled sample of HRS and PSID): retirement ages, R (baseline and alternative definitions), Social Security claiming ages, S, and life expectancy at age 60, T.

alternative. The baseline follows the definition of the consolidated labor force status in the RAND HRS: retirement is recorded when an individual is observed in a non-retired status followed by a permanent switch to the retired status.[17] The alternative definition follows Guvenen (2009): retirement is recorded when a worker's observed annual hours fall below 520 permanently using hours reported in the PSID and the HRS. Figure 1 shows that the retirement ages mildly decrease over much of the lifetime earnings distribution, only slightly increasing over the top two deciles; the Social Security claiming ages are essentially flat to a first approximation around the average of 63.5.

4.2 Productivity-age profiles

The productivity-age profiles are estimated first by adapting a standard parametric approach to our environment (e.g., Altig et al. (2001)). Let

$$\varphi(t,\theta) = \theta\varphi(t)\theta^{at},$$

where $\varphi(t)$ is a common component in age, a is a constant, and taking logarithm of both sides,

$$\log\varphi(t,\theta) = \log\theta + \log\varphi(t) + at\log\theta,$$

[17]RAND HRS reconciles all available relevant responses in each wave. In particular, it aims at separating retirement from unemployment, from partial retirement, and from reporting retirement while also reporting labor earnings.

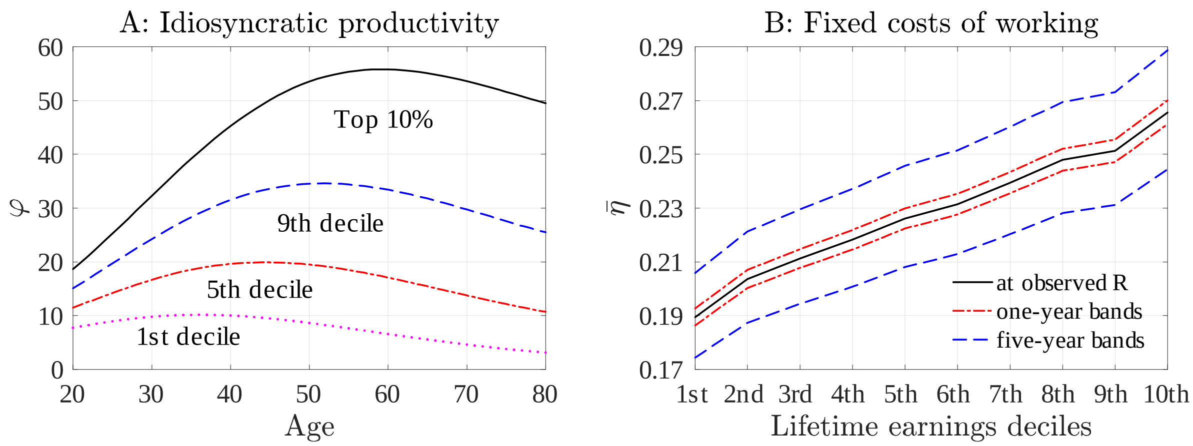

Figure 2: Idiosyncratic productivity-age profiles estimated on the pooled sample, baseline case, selected percentiles (Panel A). Calibrated fixed costs of working normalized as fractions of time endowment, at the time of observed retirement (baseline R, solid line), one year before and after observed retirement (dot-dashed lines), and five years before and after (dashed lines) (Panel B).

with the first term on the right, $\log\theta$, capturing the unobserved idiosyncratic type; the second term, $\log\varphi(t)$, is the common age profile; $at\log\theta$ captures interactions between type and age reflecting that individuals potentially age differently. Following Low and Pistaferri (2015) the individual fixed effects are interpreted as individual type, also helping address potential selection bias. To proxy for $\log\varphi(t,\theta)$ the logarithm of effective labor earnings per hour is used, i.e., the computed ratio of all labor earnings to total hours reported, converting to constant 2000 dollars (the year an individual born in 1940 would turn 60).[18] We then use predicted individual fixed effects to identify individuals into N types, producing N productivity-age profiles.

We focus here on baseline $N = 10$ with each group representing a lifetime average annual earnings decile and later vary N. Panel A of Figure 2 displays the productivity-age profiles for selected deciles, consistent with the general shape and life-cycle evolution of the profiles in the literature (e.g., Altig et al. (2001)). Higher deciles display higher productivities, generally increasing faster at younger ages. While in this general setup we no longer impose Assumption 1, its condition (i) is satisfied whenever $\varphi(t)$ decreases at older ages faster than θ^{at} increases. As a result, some fanning out is apparent in productivities with age, at least in the top deciles.

[18]The details of the generalized method of moments estimation of the resulting nonlinear statistical model are reported in the online Appendix C. A significantly more involved alternative is to construct productivities implied by the data and the individual first-order conditions, requiring a structural estimation well outside of the scope of the paper. The main challenge in that case is to correctly account for private assets that appear in the individual optimality conditions with income effects on preferences.

As with any parametric procedure, a number of concerns are debated in the literature. To address this, we show that normative findings below are robust with respect to key variations. One potential concern could be data selection: since the time variation can only be identified from the individuals who are still working, it may cause one to over- or under-estimate how fast higher productivities decline with age. Following Kahn and Lange (2014), we check robustness to changes in the curvature of the profiles, especially at older ages. We show that the effects of overestimation bias are quantitatively minimal while underestimation works to strengthen our results. Another potential concern could be the sensitivity of the data fit with respect to the parametric assumption about age dependence. To check this robustness we follow Nishiyama and Smetters (2007) grouping individual observations into bins and 10-year intervals of ages, extrapolating by using shape-preserving splines to obtain the complete productivity-age profiles. Then, we also replace age with two alternative definitions of experience to arrive at quantitatively virtually indistinguishable normative insights. The details of these as well as other robustness checks are provided in the online Appendix C.

4.3 Status-quo policies

The policy functions $\mathcal{T}$ and b are estimated to match the status-quo U.S. policies. The income tax function is given by

$$\mathcal{T}(y(t)) = y(t) - \lambda y(t)^{1-\tau}.$$

Functions of this form have been shown to approximate well the effective income taxes in the U.S., inclusive of state income taxes and a number of government non-retirement programs among others (Heathcote, Storesletten, and Violante (2014)). We follow them setting $\tau = 0.151$ (the parameter controlling progressivity) and calibrate λ to equate the present values of lifetime consumption and earnings for the cohort, which is $\lambda = 0.8067$ in our baseline. Panel A of Figure 3 shows the resulting marginal and average taxes as functions of annual earnings in constant 2000 dollars.

An intuition behind the calibration of λ is easiest to see from the overlapping-generations version of the setup, where the difference between the present values of labor earnings and consumption for a generation is equal to the total net capital income less the present value of government purchases. The net capital income is approximately payments to capital less depreciation, and in standard calibrations in the literature this net capital income as a fraction of GDP would be about $0.4 - 0.06 \times 3.5 = 0.19$, with 0.4 share of capital, 0.06 annual depreciation rate, and the capital-output ratio of 3.5, giving a historical aver-

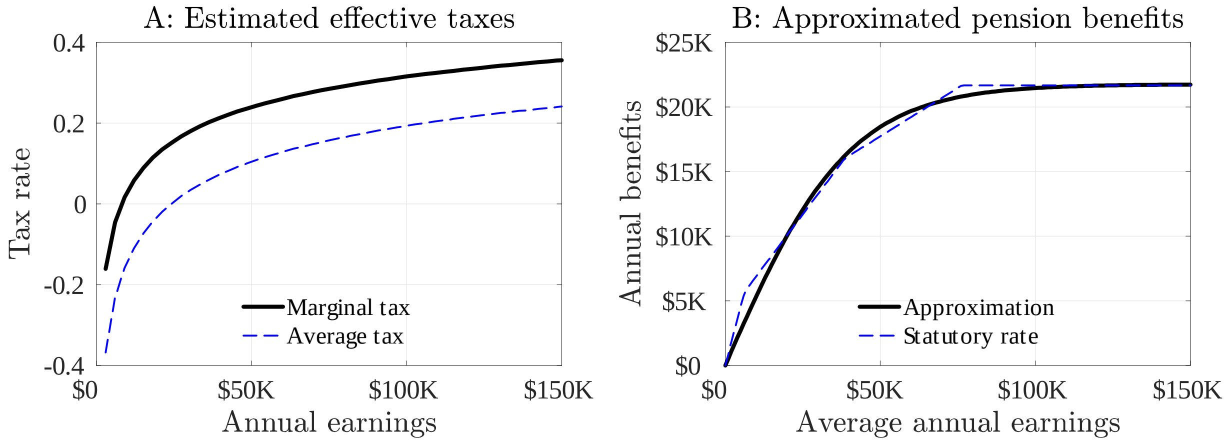

Figure 3: Estimated U.S. effective personal income tax function, $\mathcal{T}$ (Panel A). Approximated U.S. Social Security pension benefit function, $\mathcal{B}$, and annualized PIA benefits as a function of annualized AIME (Panel B). Sources: Heathcote, Storesletten, and Violante (2014), SSA (2014).

age of government purchases around 20% of GDP.[19]

To estimate the present value of net pension benefits b, we set the annual benefit function to

$$\mathcal{B}(Y) = A_1 + A_2 / \left(1 + e^{-A_3 Y}\right),$$

where Y is the average value of the highest 35 years of earnings so $\mathcal{B}(Y)$ is the annual benefit given by $b(Y) \equiv \mathcal{B}(Y) \int_S^{\bar{T}} e^{-rt} dt$, i.e., paid out between S and $\bar{T}$.[20] Parameters A_1, A_2, and A_3 are estimated by minimizing the least absolute deviation of $\mathcal{B}$ from the annualized Primary Insurance Amount (PIA) benefit formula of the Old Age, Survivor, and Disability Insurance (OASDI) part of the Social Security in 2000. The benefits are thus a function of a measure of lifetime earnings analogous to the way the PIA is a function of the Average Indexed Monthly Earnings (AIME).

Panel B in Figure 3 shows annualized $\mathcal{B}(Y)$ and for comparison PIA, as functions of annual earnings. An advantage of the estimated policies, both the tax and benefit functions, is that they capture key stylized features of the status-quo U.S. policies with relatively simple smooth functions significantly reducing the computational burden. As with the productivity-age profiles, the online Appendix C shows robustness of the normative findings below with respect to alternative estimates of status-quo policies.

[19] According the Bureau of Labor Statistics the average value of government consumption expenditure and gross investment as a fraction of GDP between 1947 and 2013 is 20.83. Nevertheless, to explore robustness we also vary this target in the next section.

[20] Recall from Figure 1 that in the U.S. micro data individual ages of claiming benefits, S, generally appear quite different from retirement ages, R.

4.4 Fixed costs

The fixed costs of working are calibrated within the structure of the positive setup. Preferences are iso-elastic with

$$u(c) = \frac{c^{1-1/\sigma} - 1}{1 - 1/\sigma}, \quad v(l) = \psi \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon},$$

where ψ is calibrated at the same time as fixed costs to match the average hours in the model to the average hours in the pooled sample. The baseline is $\sigma = 1$ with the Frisch elasticity of labor supply $\varepsilon = 0.5$ (mid-range estimates in Chetty (2012)), compared throughout with $\varepsilon = 0.25$. The interest rate is set to 2 percent with $\beta = 0.9804$.

The fixed costs are calibrated two alternative ways. First, they are simply chosen to match individually-optimal retirement ages in the model to the data. The identifying assumption in this case is that the fixed cost for a given individual type does not change at the ages around retirement, $\eta(t, \theta) = \eta(\theta)$ for t in the neighborhood of $R(\theta)$. This is weaker than condition (ii) in Assumption 1 while still providing an intuitive benchmark.

To do this, the positive setup is numerically solved for fixed costs together with allocations, while setting retirement ages, productivity-age profiles, and status-quo policies as described above. The results are shown in Figure 2 for easy comparison with productivities. The solid line in Panel B of Figure 2 plots calibrated fixed costs – at the retirement ages observed in the data – normalized as a fraction of time endowment to facilitate intuition and to allow comparisons with estimates in the literature.[21] Average fixed cost in Figure 2 is 0.2306 (e.g., Chang et al. (2014) estimate 0.2099). An interpretation is that on average fixed costs of working represent daily about five and a half hours (23.06 percent of 24 hours) – the time equivalent of commute cost and time, work attire costs, cost of meals, networking, etc. The costs increase with lifetime earnings moderately, from 0.1870 for the bottom earnings decile to 0.2659 at the top of the distribution, consistent with the estimates in the literature.[22]

Age-varying fixed costs. The second calibration of fixed costs relaxes the restriction of the costs not changing over time in order to capture other forces that may affect retirement, such as variations in health over time. A natural way to calibrate the rate of change

[21] Recall from the qualitative analysis that individual optimality requires $(1-\tau)\,\varphi_t l_t u'(c_t) = v(l_t) + \eta$ at $t = R$. The normalization means a time cost $\bar{l}$, implying $y_t = \varphi_t \max\{0, l_t - \bar{l}\}$ and the optimality condition becoming $(1-\tau)\,\varphi_t u'(c_t) = v'(\bar{l})$ at $t = R$. The time cost $\bar{l}$ is then equivalent to η if $v'(\bar{l}) = (v(\bar{l}) + \eta)/\bar{l}$ or $\bar{l} = (\eta(1+\varepsilon)/\psi)^{\frac{\varepsilon}{1+\varepsilon}}$.

[22] For a review of the estimates in the literature see, e.g., Rogerson and Wallenius (2013). For example, see $\bar{h}$ in Table 1 in Chang et al. (2014) with $\gamma = 0.5 = 1/\varepsilon$ and the lowest σ_x, i.e., closest to permanent shocks (their $B = 82.70$ is the counterpart of our calibrated $\psi = 84.66$).

in fixed costs with age is to target existing estimates of the elasticity of retirement age with respect to policies affecting retirement decisions. The literature successfully identifying such responses empirically is limited (Chetty et al. (2011)).[23] We instead follow French (2005) and use the estimates from his structural life-cycle model of labor supply, health, and retirement. French (2005) reports a range 1.04 - 1.33 of estimates of the elasticity of labor supply with respect to a temporary anticipated increase in earnings at age 60 (see, e.g., his Table 2). He shows that at those ages the individuals are close enough to retirement so that the labor response is coming mainly from the participation decision at the retirement margin.

We numerically re-solve the positive setup choosing the rate of change in fixed costs to match the elasticity, while setting fixed costs at the observed retirement ages to the point estimates above since a temporary anticipated increase in earnings leaves unaffected the point-identification. In other words, in the complete calibrated positive setup the fixed costs at the observed retirement ages are pinned down by the individual optimality conditions at those ages while the rate of change of fixed costs with age is calibrated to match the average labor supply elasticity at age 60. The rate of change is calibrated to 1.68% per year targeting the elasticity 1.185, the middle of the range in French (2005). The resulting evolution of fixed costs with age is illustrated also in Panel B of Figure 2. The dot-dashed lines show 1-year bands around the observed retirement ages: fixed costs 1 year before observed retirement ages are shown as a dot-dashed line below the solid line of point estimates; fixed costs 1 year after observed retirement are shown as a dot-dashed line above the point estimates. Similarly, the 5-year bands around the observed retirement ages are shown by the dashed lines. On average, fixed costs 5 years after the retirement ages observed in the data are the equivalent of just over 6 hours per day and by age T the costs exceed 8 hours per day.

Given the robustness checks on the productivity-age profiles and status-quo policies above, to be consistent verifying robustness of normative findings in the next section we re-calibrate fixed costs for each of the robustness checks. As a check of external validity of the positive model we also simulate generic extensive labor supply elasticity: starting from the baseline with retirement ages matching the data we compute required simulated change in the present value of lifetime pension benefits to increase the average retirement

[23] The robustness of existing estimates is limited primarily by data availability on older individuals' working behavior over time, resulting in wide ranges of estimates and challenges singling out the exact elasticity being estimated. A particular challenge is disentangling income and substitution effects. The changes in wages arguably have both income and substitution effects and could be endogenous to health and therefore retirement decision. At the same time, there is limited variation in pension benefits among groups of similar individuals near retirement age.

age by one year, giving the average elasticity of 0.37 within the range $0.13 - 0.43$ in the literature (Chetty et al. (2011)). Another standard external validity check is to compare simulated stationary income distribution against an alternative data source. According to the U.S. Census, the median male's income in the U.S. in 2000 was $37,339 (with a 90% confidence interval of ± 225.0), while our baseline calibration here is $34,285. The bottom 10 percent of the distribution had average income of $15,722 ($\pm 94.7$) vs. calibrated $16,515; the top decile was $85,487 ($\pm 515.1$) vs. calibrated $89,905.

5 Quantitative Analysis of Constrained Optima

Given the estimated parameters of the general environment, we return to normative questions to quantify the insights identified by the analysis in Section 3 as well as what they would imply for the overall welfare and aggregate output. Taking advantage of the qualitative insights it is tractable to numerically solve for the constrained-efficient allocations over the complete life-cycle, verify ex post global incentive constraints, and simulate optimal policy implementations. We start with the baseline estimates of productivity-age profiles and fixed costs and with exogenous redistributive motives given by equal weights in the social welfare function, $G(\theta) = F(\theta)$ for all $\theta \in \Theta$. We then vary the welfare weights and the rest of the parameters as described in the previous section to explore quantitative robustness and extensions. Broad quantitative insights that emerge from comparing constrained optima to the status quo are (i) significantly later retirement of higher earning individuals, (ii) increased benefits from delayed retirement, which further increase with lifetime earnings, and (iii) large welfare and output gains from switching to optimal policies.

5.1 Efficient retirement ages

Proposition 2 suggests that the profile of constrained-efficient retirement ages across types can be captured in a straightforward way by the differences in virtual productivities relative to fixed costs. Figure 4 illustrates these relative differences as a function of lifetime earnings, evaluated at three sets of ages: at the retirement ages observed in the U.S. microeconomic data, R, at the observed Social Security PIA-benefit claiming ages, S, and at an arbitrary constant age 70. The expression evaluated at the observed retirement ages, R, is plotted with a solid line. Given that it is increasing almost everywhere with lifetime earnings, Proposition 2 implies that retirement ages should also be increasing almost everywhere to be constrained-efficient, which is in contrast with almost flat profiles found in

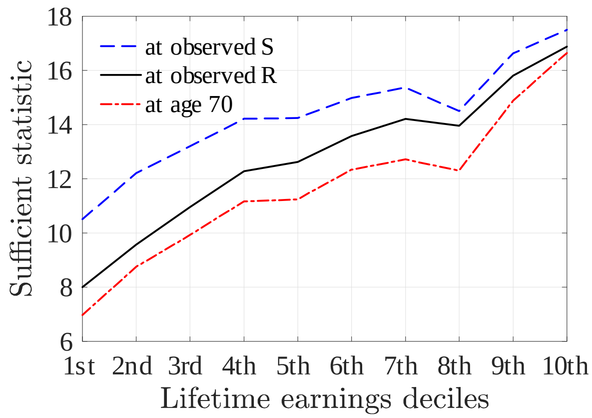

Figure 4: Sufficient statistic expressions for retirement ages, $\tilde{\varphi}(t,\theta)^{1+\varepsilon}/\tilde{\eta}(\theta)$, evaluated at the retirement ages observed in the U.S. microeconomic data, R (baseline definition), at the Social Security claiming ages, S, and at a constant age 70.

Figure 1 for the U.S. data.

This type of sufficient statistic expression allows a straightforward evaluation of counterfactual experiments often put forth as potential goals for pension reforms. We first consider a counterfactual where every worker in the U.S. instead is required to retire when they claim their Social Security benefits, i.e. at ages S. Evaluating the expression at S for each type produces the dashed line in Figure 4 which is once again increasing for almost all types, suggesting once again that a flat retirement age profile would be suboptimal. Illustrated with the dot-dashed line in Figure 4, the same conclusions are reached when we consider a counterfactual with every worker retiring instead at a constant age, for instance, 70. What these experiments suggest is that reforms that do not correctly account for the *individualized* incentives relevant for retirement decisions are likely to keep the economy away from constrained-efficiency.

To break down the sources of increasing expressions above recall that we have already seen above in Figure 2 that both the productivities and the fixed costs generally increase with type at a given age. Comparing vertical scales in Panels A and B in Figure 2, it is also apparent that the productivities change much faster with type, in fact close to exponentially, while fixed costs change only about linearly. This is quite intuitive since one expects workers to be potentially vastly more different at productive tasks rather than at commute, meals, attire, etc. In terms of virtual productivities and fixed costs, Proposition 1 further implies that information rents only strengthen this relationship (except in the presence of extreme redistributive motives, e.g., Rawlsian motives that we consider below).

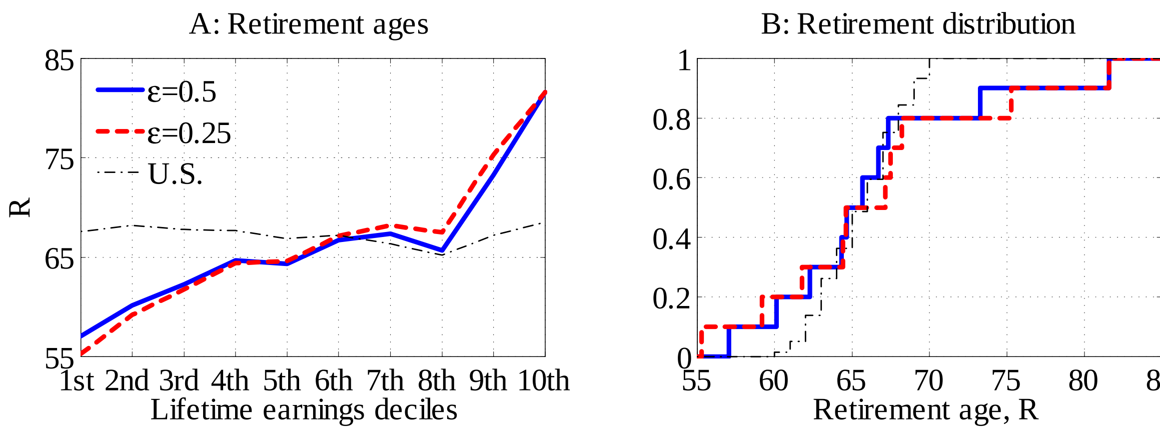

Figure 5: Constrained-efficient retirement ages: As a function of lifetime earnings with Frisch elasticity $\varepsilon = 0.5$ (solid line) and with $\varepsilon = 0.25$ (dashed line) compared to U.S. microeconomic data (R, baseline definition, dot-dashed line) (Panel A). Cumulative distributions of retirement ages (Panel B).

Turning to the constrained-efficient allocations themselves, Figure 5 illustrates quantitative observations that are robust across variations in the estimated parameters. The solid line in Panel A plots constrained-efficient retirement ages as a function of lifetime earnings in the baseline case with Frisch elasticity $\varepsilon = 0.5$,[24] contrasted with the dot-dashed line plotting the retirement ages observed in the U.S. micro data (dashed line provides a comparison with $\varepsilon = 0.25$). In line with the analysis of Figure 4, efficient retirement ages increase for almost all lifetime earnings and strictly increase in the tails of the distribution. The retirement starts from age 57.1 at the bottom of the distribution increasing to 60.1 by the second decile and to 64.5 by median lifetime earnings. The top three deciles show a further increase, bringing the average retirement age for the top half of the distribution to 69.8. The variation in the Frisch intensive elasticity only slightly alters the quantitative extent of these increases. Intuitively, lower intensive elasticity moderately exacerbates individual differences in productivities relative to fixed costs, which would make profiles in Figure 4 somewhat steeper, hence implying a somewhat steeper retirement age profile.

The constrained-efficient profiles of retirement ages in Figure 5 are also notably steeper than in the U.S. microeconomic data. The mechanics behind this are that in the bottom half of the distribution the expressions at the observed retirement ages in Figure 4 are lower than the constrained-efficient value (which would be just below 14), suggesting

[24] To facilitate comparisons, this baseline case is plotted with a thick solid line throughout the quantitative analysis as well as later throughout the extensions.

that those are the ages that are suboptimally high in the data. The opposite holds for the top 20 percent of the distribution (while the 6th through 8th deciles are close to the efficient levels). A comparison of the constrained-efficient solid-line profile with the dot-dashed line of the U.S. retirement ages in Panel A of Figure 5 bears out these observations. Workers in the top half of the lifetime earnings distribution efficiently retire on average at the age of 69.8, while in the U.S. data their average retirement age is 66.9. Workers in the bottom half retire at 61.7, while in the U.S. data they do so at 67.6. For an alternative comparison Panel B also maps these observations into cumulative distributions of retired individuals that are conventional in the literature. Once again the variation in the intensive elasticity minimally alters the quantitative differences.

5.2 Optimal pension systems

Consider next the differences between optimal and the status-quo policies. Another perspective on the sufficient statistic expressions is that the above analysis implies that in the constrained optimum more productive workers need to be provided with incentives for later retirement. A consequence of the expressions in Figure 4 being above the constrained-efficient values at the top of the lifetime earnings distribution is that the status-quo policy distorts retirement too much in that part of the distribution, resulting in retirement too early. Similarly at the bottom of the distribution the distortion is also too strong (in absolute value) albeit distorting retirement in the opposite direction. Since Proposition 3 shows that retirement incentives are costlier than incentives for the hours decisions, this implies that the status-quo policies cannot be optimal.

Wedges. We simulate first the implied wedges in order to illustrate optimal incentives independently of the particulars of an implementation. Panel A in Figure 6 plots the retirement wedge, τ_R, as a function of lifetime earnings with a solid line for the baseline $\varepsilon = 0.5$ (as before, the dashed line provides a comparison with $\varepsilon = 0.25$). The positive quantitative framework allows us to also simulate the retirement wedges implied by the U.S. data in the previous section. We plot them as a function of lifetime earnings with a dot-dashed line in Panel A. As explained by the above analysis, for most of the earnings levels and especially in the tails the retirement distortions are reduced in absolute size in the optimum versus the distortions by the status-quo policies. A takeaway is that the existing U.S. policies imply retirement wedges that are significantly more progressive than optimal. That is, quantitatively they distort retirement age decisions in the tails of the earnings distribution too much given how costly the retirement incentives are relative to the hours incentives.

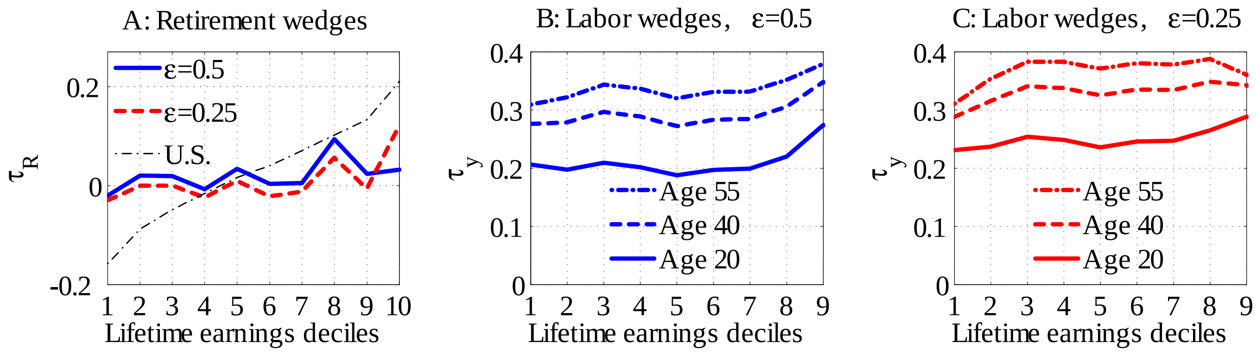

Figure 6: Optimal wedges: Retirement wedge, τ_R, with baseline $\varepsilon = 0.5$ and with $\varepsilon = 0.25$ compared with the retirement wedge implied by the status quo U.S. policies (Panel A). Labor wedges, $\tau_y(t)$, at selected ages with baseline $\varepsilon = 0.5$ (Panel B) compared with $\varepsilon = 0.25$ (Panel C).

The labor wedges, τ_y, are shown as functions of lifetime earnings in Panels B and C of Figure 6 for selected ages: at the beginning of the working life (age 20, solid line), at prime (age 40, dashed line), and closer to retirement (age 55, dot-dashed lines). Intuitively, since relative differences between individual productivities tend to increase with age, so do the labor wedges to provide incentives to work at full productivity as explained by our generalization of the standard marginal tax formula (12). The lower Frisch elasticity leads to higher wedges as well. A quantitatively key insight from Figure 6 is that the retirement wedge is of smaller value than the labor wedges everywhere. Recall that our analysis of Propositions 3 and 4 shows that in this case optimal pension systems reward later retirement, regardless of whether the efficient retirement ages are increasing or decreasing. The quantitative differences between the wedges are almost an order of magnitude, implying that sizeable incentives for later retirement are needed to offset the incentives for earlier retirement imbedded in marginal income taxes.[25]

Policies. Turning to optimal policy functions, recall from Lemmas 1 and 2 that implementation characterizes marginal policies with the total functions identified up to a

[25]Note also that while standard results in the literature may suggest to attribute increasing efficient retirement ages to the standard optimality of no distortions for the top type, it cannot be the explanation here. For the top decile, the undistorted labor supply produces a large increase in retirement age from the status quo 68.5 to 81.6, as we have seen in Figure 5. The increases in retirement age, however, are not limited to the top decile. The labor wedge evaluated at retirement for the 9th decile is around 40 percent, yet their retirement age also significantly increases compared with status quo, from 67.2 to 73.6 . Figure 6 reveals that this is due to a significant reduction in their retirement wedge compared with the status quo. Analogous observations hold for the top half of the distribution with the opposite observation for the bottom tail, while the middle of the earnings distribution in the data experience close to optimal retirement incentives as we showed above.

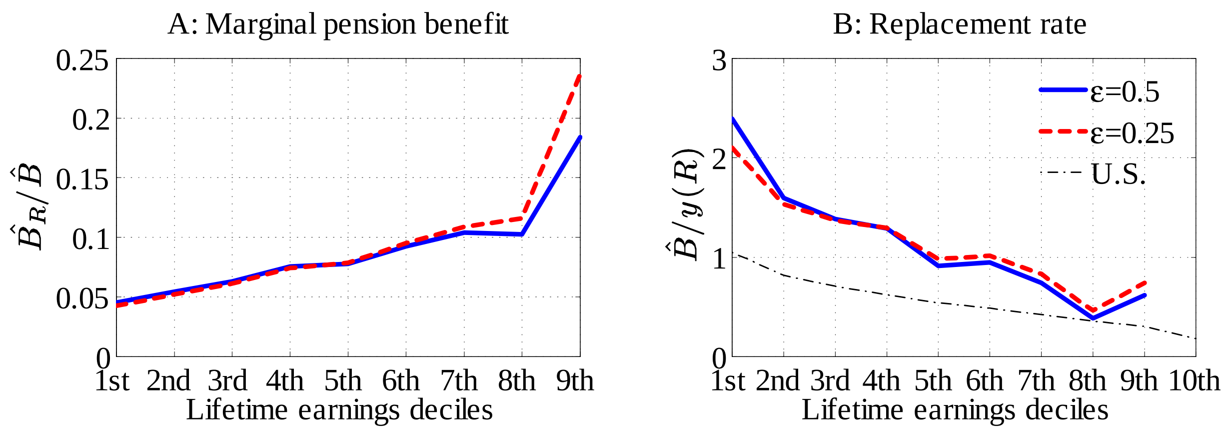

Figure 7: Optimal pension benefits encourage later retirement (Panel A, marginal benefits with respect to an increase in retirement age). Optimal replacement rates compared to the replacement rates implied by the U.S. microeconomic data and the status-quo U.S. Social Security (Panel B).

constant. To fix the constant (the intercept of the income tax function) we keep the average tax for the bottom decile at -33.3 percent as in the status quo.[26] Then the optimal income tax system is fully quantified with the marginal tax rates given by the above analyzed labor wedges in Figure 6. The optimal pension benefits are illustrated in Figure 7. Panel A plots marginal (with respect to an infinitesimal increase in retirement age) pension benefits as a function of lifetime earnings, $\hat{\mathcal{B}}'(R)/\hat{\mathcal{B}}(R)$. The two lines represent the baseline elasticity $\varepsilon = 0.5$ (solid line) and $\varepsilon = 0.25$ (dashed line). Recall that for a type θ facing income taxes $\mathcal{T}(t, y)$ and a per-period pension benefit $\hat{\mathcal{B}}(R)$ the optimality of the choice of R requires

$$v\left(\frac{y(R,\theta)}{\varphi(R,\theta)}\right) + \eta(\theta) =$$
$$u'(c(\theta))\, y(R,\theta)\left(1 - \frac{\mathcal{T}(R, y(R,\theta))}{y(R,\theta)} - \frac{\hat{\mathcal{B}}(R)}{y(R,\theta)} + \hat{\mathcal{B}}'(R)\frac{1-e^{-r(\bar{T}-R)}}{r}\right)$$

where intuitively $\hat{\mathcal{B}}'(R)$ is the rate of return on choosing to retire a bit later so that $\hat{\mathcal{B}}'(R)/\hat{\mathcal{B}}(R)$ implements the retirement wedge $\tau_R(\theta)$.[27]

Panel A shows that the optimal pension system rewards later retirement as explained

[26] According to CBO (2012) the disposable income including transfers of a single parent with one child with an income of \$15,000 is approximately \$20,000, which amounts to an average tax rate of -33.3 percent. Alternatively, the constant can be fixed, for example, to the rate for the median individual without qualitatively changing the results.

[27] This illustrates the central mechanism and since the lifetime earnings could also depend on the retirement age the online Appendix B provides the details of computing a complete implementation.

by the analysis of Proposition 4. The optimal return on later retirement is positive throughout the distribution, ranging from 4.5% at the bottom to 18% at the top of the distribution in the baseline. Optimal marginal benefits increase with earnings, especially in the right tail of the distribution. Intuitively, lower elasticity requires a somewhat steeper profile to provide the same incentives. Our quantitative analysis here also allows a direct comparison with the status-quo U.S. policies. According to SSA (2014), Table 2.A20, the rate of return on later retirement ranged $3.5\% - 8\%$ between 1987 and 2005. While this is in the same direction as the optimal policy, the quantitative magnitudes are substantially different. Moreover, the status-quo rate of return on later retirement was flat throughout the distribution while it sharply increases in the optimal policy in Panel A of Figure 7.

Panel B further quantifies the comparison with the status quo by simulating replacement rates, i.e., the ratio of annual pension benefits to the average annual earnings during the working life, which is a common measure in the literature.[28] As before, the solid line plots the baseline $\varepsilon = 0.5$ replacement rates, the dashed line plots replacements rates with $\varepsilon = 0.25$, and the dot-dashed line provides a comparison with the U.S. status-quo replacements rates. The optimal replacement rates decrease as a function of lifetime earnings and in the right tail of the distribution reach levels similar to the status quo. At the bottom of the distribution, relatively earlier constrained-efficient retirement ages and negative average income taxes at retirement require the optimal replacement rate of almost 240 percent compared with just over 100 percent in the status-quo U.S. Social Security system. In other words, as above, the takeaway is a quantitatively significant increase in returns to later retirement.

5.3 Welfare and aggregate output effects

To quantify the importance of the differences illustrated above, we compute the changes in welfare via equivalent consumption variation as well as the changes in output from switching from the U.S. status quo to the baseline simulation of the constrained efficiency.

Table 2 compares both the aggregate values and selected percentiles of the lifetime earnings distribution. The top part of Table 2 shows that the aggregate welfare change for the cohort is 7.1% in consumption equivalent. As the analysis above explains, the welfare in the right tail of the distribution declines – by as much as 20% in the top ten percent of the distribution – while the bottom of the distribution gains significantly – the average of about 12% in the bottom half of the distribution.

[28]Note that the replacement rate is with respect to earnings (not contributions) and will have no effect on the capital stock here. In fact, the pension system does not even have to be pay-as-you-go.

Table 2: Effects of switching to baseline optimum from the status quo.

Changes in Welfare				
Aggregate	1st decile	2nd decile	9th decile	10th decile
7.1%	44.9%	24.4%	-4.7%	-20.0%

Changes in Output				
Aggregate	1st decile	2nd decile	9th decile	10th decile
1.7%	-26.0%	-18.7%	1.8%	35.0%

To understand the sources of large changes in welfare throughout the earnings distribution as well as of the gains in the aggregate, consider how the constrained efficiency affects the output of the economy when the pension system provides individuals with optimal incentives for retirement. The bottom part of Table 2 reveals a substantial growth in aggregate output, increasing it by 1.7 percent. The main driving force behind the increase in aggregate output is the increase in retirement ages we saw in the right tail of the distribution. The combination of reduced retirement wedges we showed above and the persistence of productivity differences with age imply large output gains at the top of the earnings distribution – by more than a third for the most productive decile. Quantitatively, those output gains are more than enough to offset the losses needed for the earlier retirement of the workers in the left tail of the distribution. The resulting significant aggregate output gains contrast with most of the literature, indicative of the potentially underexplored power of retirement incentives as an optimal policy focus.

5.4 Other forces affecting retirement

Next we explore other mechanisms that could reasonably be expected to potentially affect quantitative findings. We focus here on key mechanisms while the online Appendix provides further details of the robustness of the quantitative insights with respect to the alternative definitions of retirement age, alternative estimates of the productivity-age profiles, accounting for alternative measures of experience, varying the steepness of the productivity decline at older ages, alternative status-quo policy approximations, and the number of types. As before, we illustrate the effects focusing on constrained-efficient retirement ages and optimal pension systems to summarize main insights.

Heterogeneous life span. Given considerable evidence that mortality is closely related to lifetime earnings (Waldron (2013) among others), we allow for heterogeneous life span in order to explore the effects longer lives could have on efficient retirement and optimal policies by changing the total amount of lifetime benefits collected. This allows us to keep the focus on the efficient retirement ages and optimal pension systems while abstracting from numerous related issues beyond the scope of the paper.[29]

The environment is as given in Section 2 except the length of life for each individual of type θ is now given by $T(\theta)$ where $T : \Theta \to \mathbb{R}_+$ is a typically increasing function. The utility function for each individual of type θ is then given by

$$\int_0^{T(\theta)} e^{-rt} u\left(c\left(t,\theta\right)\right) dt - \int_0^{R(\theta)} e^{-rt} \left[v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)} \right) + \eta\left(t,\theta\right) \right] dt. \tag{15}$$

The constrained-efficient allocation is defined as before to be the solution to the mechanism design problem except with preferences given by (15). We re-state the complete modified mechanism design problem in the online Appendix B and formally show that the main channel through which life span can potentially affect the retirement decision is through its income effect on consumption. Thus the quantitative environment can be constructed analogously to Section 4. To obtain the extension of $\bar{T}$ to $T(\theta)$ we follow Waldron (2013). We use death rates from Chart 1 in Waldron (2013) and apply a standard lifetable calculation. In particular, to extend the death rates beyond the reported age of 71, we calibrate the growth rate in death rate to match the average life expectancy of the cohort to that reported in Bell and Miller (2005) for the 1940 birth year cohort males, given they survived to age 60. The growth rate is calibrated to be 0.37 to match the average cohort life expectancy of 81.6. As a secondary check, we compare the implication for the difference in life expectancy between the lower and the upper halves of the income distribution to those given in Chart 3 in Waldron (2013). For easier comparison earlier in Figure 1 the dot-dashed line plots the average life expectancy for each lifetime earnings decile.

Figure 8 plots the resulting efficient retirement ages when life span is heterogeneous (thick dashed line in Panel A) together with the baseline efficient retirement with homogeneous life spans (thick solid line in Panel A). For reference, life expectancy as a function of lifetime earnings is compared with the baseline $\bar{T}$ in Panel A as well (thin dashed and solid lines respectively). As the analysis in Section 3 suggested, the differences between the efficient retirement ages in the two cases are virtually negligible with the only exception in the top decile of lifetime earnings, who still have to be undistorted but now have a longer life-cycle. This implies that the heterogeneity in life span by itself does not produce large income effects. This however does not mean that the structure of the pension

[29] One such aspect of heterogeneous life spans is mortality risk, which adds demand for insurance and makes annuities important. See, e.g., Hosseini and Shourideh (2016), who study the effects of differential mortality on optimal taxes while abstracting from the retirement decision.

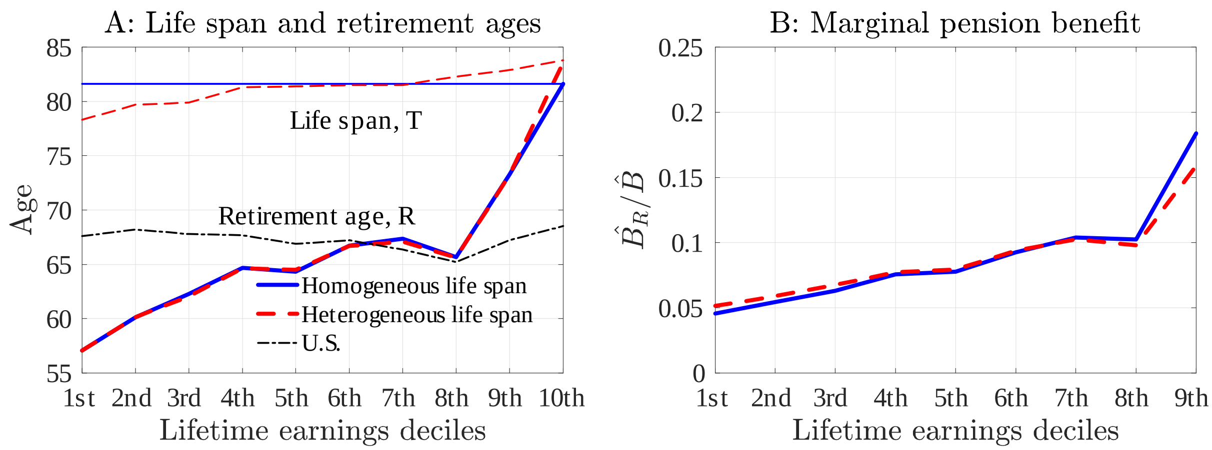

Figure 8: The effects of heterogeneity in the life span on efficient retirement ages (Panel A) and on optimal pension design (Panel B).

benefits has to stay exactly the same. In fact, the heterogeneity in life span mostly affects the structure of pension benefits and not that of earnings. The quantitative differences in the retirement wedges are still small and hence Panel B illustrates this point with marginal pension benefits. The heterogeneity in life span results in a somewhat flatter profile of marginal pension benefits (dashed vs. solid lines in Panel B), closely resembling to the effects of higher intensive elasticity analyzed above in Figure 7.

Intergenerational transfers. A potential limitation is the assumption that the present value of earnings is equal to the present value of consumption for the cohort, assuming away inter-generational transfers through the pension system. An existing pay-as-you-go Social Security system, however, may tend to decrease the present value of consumption relative to that of earnings depending on the replacement rates for a particular generation. Our baseline assumption is not obviously innocuous because intergenerational transfers can affect the labor supply decisions about both the hours worked and the retirement age through income effects.

To explore the effects of this argument on our quantitative findings, we next allow for the variations in the difference between the present values of consumption and earnings. Specifically, we modify the calibration of parameter λ in the tax function we described above so that this difference is 10 percent (i.e., net transferring) and also -10 percent (i.e., net receiving) of the present value of earnings. Figure 9 illustrates the resulting changes to the efficient retirement ages in Panel A and to the optimal retirement wedges in Panel B to summarize the optimal pension system. Expectedly, as the cohort becomes richer in the sense of receiving intergenerational transfers, income effects drive earlier retirement ages.

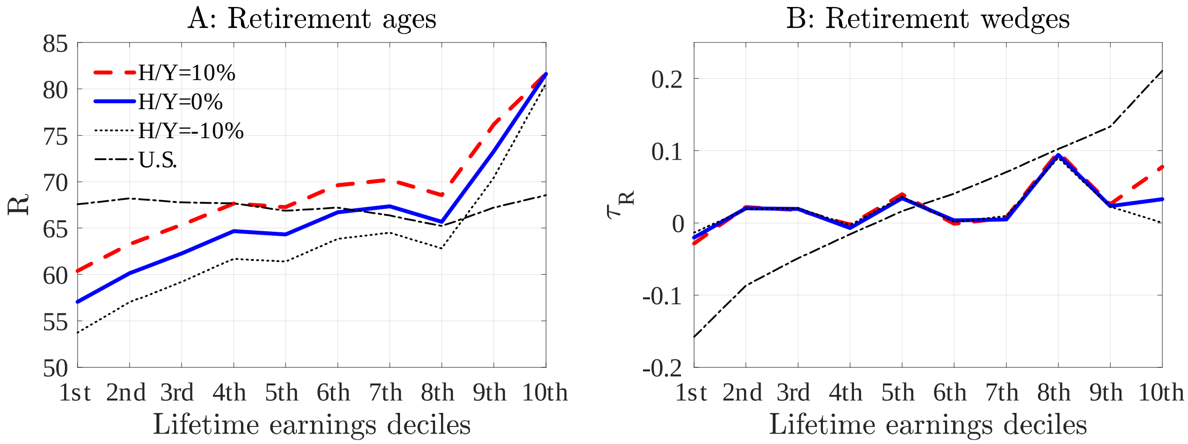

Figure 9: The effects of intergenerational tranfers on efficient retirement ages (Panel A) and on optimal pension design (Panel B).

On the other hand, as the cohort transfers part of its earnings the income effects require a steeper retirement age profile. However, these changes do not overturn the overarching insights above about the retirement age profiles. The policy implications are once again similar, with the only noticeable differences in the retirement wedge at the very top of the distribution where it is still optimal to provide incentives to work throughout the life-cycle.

Effects of age-varying fixed costs. To quantify the effects of the fixed costs varying with age, we compare the constrained optima using the two alternative calibrations of fixed costs discussed in the previous section. The results are illustrated in Figure 10. Panel A plots the resulting efficient retirement ages with age-varying fixed costs (dashed line) compared with efficient retirement with fixed costs that do not change around the time of retirement (solid line). Similarly, Panel B compares the optimal retirement wedges. To provide a common point of reference, the dot-dashed lines plot the retirement ages observed in the U.S. data (Panel A) and the retirement wedges implied by the data (Panel B). The general insight that more productive individuals retire later than in the U.S. data does not depend on the fixed costs varying with age. Similarly, unlike in the status quo the efficient retirement ages increase over almost all of the lifetime earnings distribution with or without the fixed costs varying with age. While tight identification of fixed costs away from observed retirement behavior is challenging as we discussed in the previous section, this suggests that general insights are quite robust to age-variations.

As our analysis in Section 3 explains, there are two competing effects: a mechanical effect of varying costs and an information effect from altered incentives. Start at the bottom

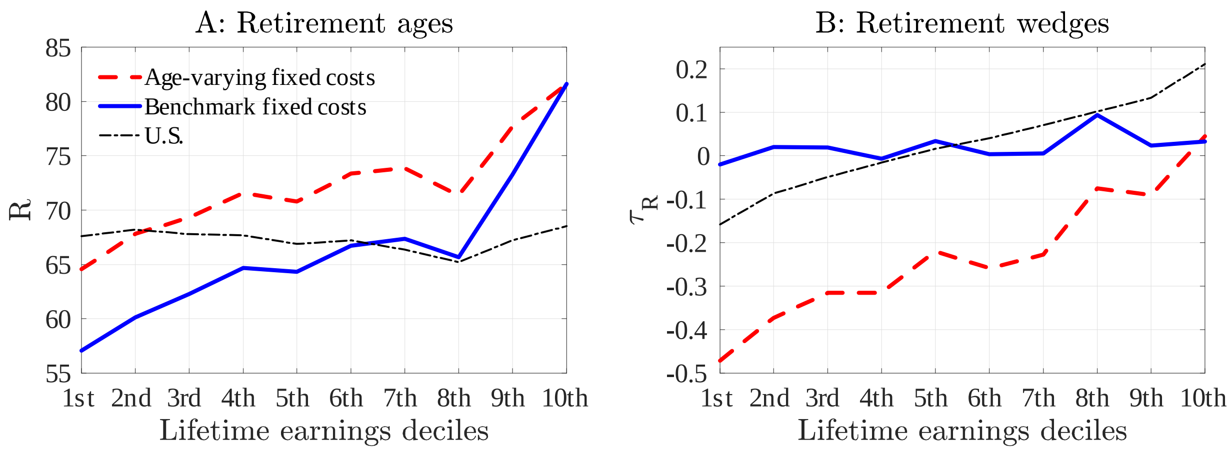

Figure 10: The effects of varying fixed costs with age on efficient retirement ages (Panel A) and on optimal pension design (Panel B).

of the lifetime earnings distribution where there are no information effects. Compared with stable fixed costs, age-variation lowers costs at the bottom of the distribution since the efficient retirement ages are below the ones observed in the data (the fixed costs in the two cases are the same at the observed retirement ages). Thus mechanically it becomes efficient for the workers at the bottom of the earnings distribution to retire later than if the costs were not varying (the dashed line is above the solid line at the bottom of the distribution in Panel A).

This later efficient retirement weakens the incentives of higher-earnings workers to misreport their types, lowering their information rents. Once a worker is high enough in the distribution, however, age-variation increases fixed costs compared with stable fixed costs since the efficient retirement ages are above the ones observed in the data. That creates a trade-off between higher fixed costs and lower information rents: on the one hand, higher fixed costs mechanically require earlier retirement; on the other hand, later retirement lower along the distribution lowers information rents, requiring later retirement. The fact that the dashed line is above the solid line in Panel A in Figure 10 everywhere except at the very top shows that the mechanical cost-variation effects are dominated by the information effects. Consequently retirement ages go up for all but the very top of the distribution.

Panel B illustrates this mechanism from the point of view of the optimal pension system. As a result of the above, there is no need for the optimal system to distort retirement age decisions as much and consequently with age-varying fixed costs the retirement wedges are even lower, and everywhere lower than the ones implied by the status quo.

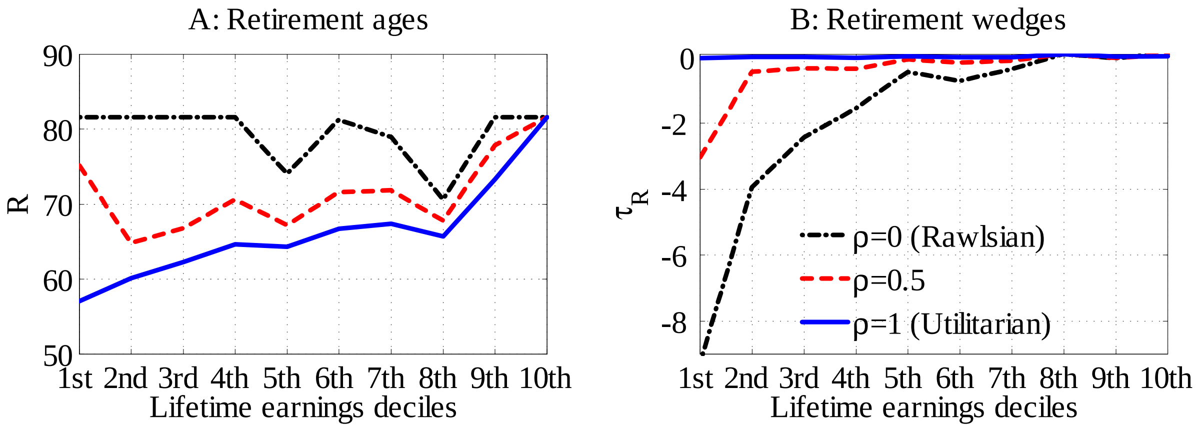

Figure 11: The effects of higher redistribution on efficient retirement ages (Panel A) and on optimal pension design (Panel B).

Alternative notions of social welfare. To explore the sensitivity of quantitative findings with respect to changing redistributive motives, we set the cumulative social welfare function $G(\theta) = F(\theta)^{\rho}$ where $\rho \geq 0$. The case of $\rho = 1$ corresponds to the Utilitarian welfare function analyzed in the baseline case above. The cases of $0 \leq \rho < 1$ are more redistributive in the usual sense with the extreme of $\rho = 0$ corresponding to the Rawlsian criterion. The cases of $\rho > 1$ are the opposite extremes of redistributing away from the bottom of the distribution.

Figures 11 and 12 illustrate the effects on retirement ages in Panels A and on the retirement wedges in Panels B. As ρ decreases toward 0, the planner tends to allocate more utility to the left tail of the distribution; in particular, as ρ approaches 0, the left tail of the distribution has to work throughout their life. Intuitively, late retirement prevents more-productive individuals from pretending to be less productive and decreasing hours while they work, which explains the necessity of large negative wedges in the left tail shown in Panel B of Figure 11. As ρ increases above 1, the level of redistribution toward the left tail declines, producing income effects leading the bottom deciles to retire later; the same redistributive mechanism as above now leads the higher deciles to lower their retirement age except at the very top of the distribution, which has to remain undistorted.

6 Conclusions

We studied the trade-offs between efficiency and incentive costs of social insurance and redistribution when retirement decisions are endogenous. Our equivalence result pro-

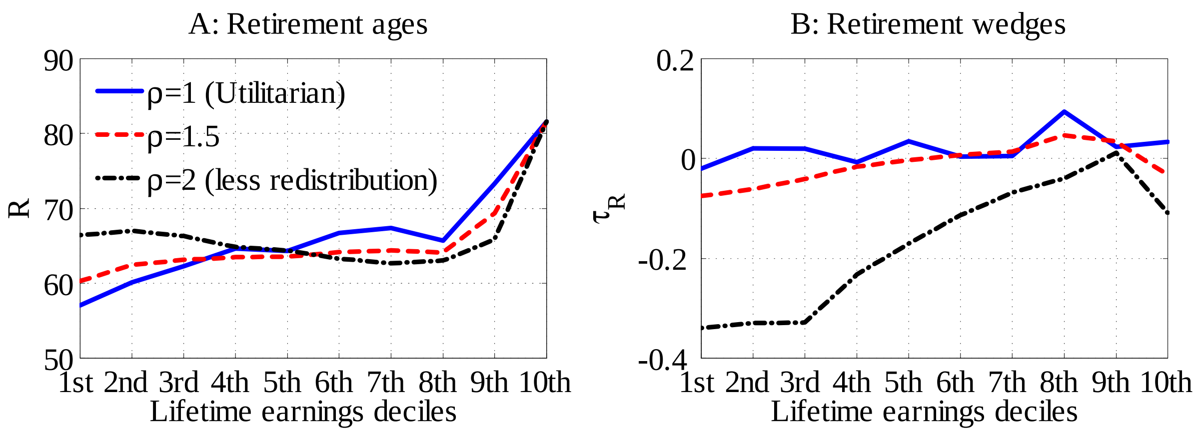

Figure 12: The effects of less redistribution on efficient retirement ages (Panel A) and on optimal pension design (Panel B).

vides a straightforward characterization of the key forces behind constrained optima and, in particular, behind constrained-efficient retirement ages. We showed that pension systems that reward later retirement are optimal independent of increasing or decreasing patterns of retirement ages. We also showed that providing correct individualized work and retirement incentives can result not only in significant aggregate welfare gains but also in aggregate output gains. We focused on permanent shocks following the findings that most of the welfare gains come not from redistribution with temporary shocks but from the provision of insurance against permanent shocks (see, e.g., Huggett and Parra (2010)). It suggests that while including temporary shocks may adjust quantitative differences between the status quo and the optimum, the qualitative changes that we characterized here would remain.[30]

References

Ábrahám, Árpád, Sebastian Koehne, and Nicola Pavoni. 2011. "On the first-order approach in principal-agent models with hidden borrowing and lending." *Journal of Economic Theory* 146 (4):1331–1361.

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters, and Jan Walliser. 2001. "Simulating Fundamental Tax Reform in the United States." *The American Economic Review* 91 (3):574–595. 7, 18, 19

[30]An extension of our framework develops this argument formally in an earlier working paper available upon request.

Bell, Felicitie C and Michael L Miller. 2005. "Life Tables for the United States Social Security Area 1900-2100." *Social Security Administration Publication* 11 (11536). 17, 32, xx

Best, Michael and Henrik Jacobsen Kleven. 2013. "Optimal Income Taxation with Career Effects of Work Effort." *Working Paper* . 7

CBO. 2012. "Effective Marginal Tax Rates for Low- and Moderate-Income Workers." *Congressional Budget Office* . 29

Chang, Yongsung, Sun-Bin Kim, Kyooho Kwon, and Richard Rogerson. 2014. "Individual and Aggregate Labor Supply in a Heterogeneous Agent Economy with Intensive and Extensive Margins." *Working Paper* . 22

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1 (1):451–488. 12

———. 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80 (3):969–1018. 22

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2011. "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." *NBER Working Paper No. 16729* . 23, 24

Choné, Philippe and Guy Laroque. 2015. "On the Redistributive Power of Pensions." *Working Paper* . 3, 8

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing capital? Not a bad idea after all!" *American Economic Review* 99 (1):25–48. 4, 7

Cremer, Helmuth, Jean-Marie Lozachmeur, and Pierre Pestieau. 2004. "Social Security, Retirement Age and Optimal Income Taxation." *Journal of Public Economics* 88 (11):2259–2281. 3, 8

Farhi, Emmanuel and Iván Werning. 2013. "Insurance and Taxation over the Life Cycle." *The Review of Economic Studies* 80 (2):596–635. 4, 7

French, Eric. 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies* 72 (2):395–427. 23

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. MIT Press, Cambridge, Massachusetts. vi

Golosov, Mikhail, Ali Shourideh, Maxim Troshkin, and Aleh Tsyvinski. 2013. "Optimal Pension Systems with Simple Instruments." *American Economic Review* 103 (3):502–07. 4

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski. 2011. "Optimal Dynamic Taxes." *NBER Working Paper No. 17642* . 4, 13

Guvenen, Fatih. 2009. "An Empirical Investigation of Labor Income Processes." *Review of Economic Dynamics* 12 (1):58–79. 17, 18, xxi

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006." *Review of Economic Dynamics* 13 (1):15–51. 17, xviii

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2014. "Optimal Tax Progressivity: An Analytical Framework." *NBER Working Paper No. 19899* . 20, 21

Hosseini, Roozbeh and Ali Shourideh. 2016. "Retirement Financing: An Optimal Reform Approach." *Working Paper* . 32

Huggett, Mark and Juan Carlos Parra. 2010. "How Well Does the U.S. Social Insurance System Provide Social Insurance?" *Journal of Political Economy* 118 (1):76–112. 4, 7, 37

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review* 101 (7):2923–54. 4, 7

Kahn, Lisa B and Fabian Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *The Review of Economic Studies* 81 (4):1575–1613. 20

Kapička, Marek. 2013. "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach." *The Review of Economic Studies* 80 (3 (284)):1027–1054. 10

Low, Hamish, Costas Meghir, and Luigi Pistaferri. 2010. "Wage Risk and Employment Risk over the Life Cycle." *American Economic Review* 100 (4):1432–67. 4, 16

Low, Hamish and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off." *American Economic Review* 105 (10):2986–3029. 4, 19

Michau, Jean-Baptiste. 2014. "Optimal Redistribution: A Life-Cycle Perspective." *Journal of Public Economics* 111:1–16. 3

Myerson, Roger B. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6 (1):58–73. 11

Nishiyama, Shinichi and Kent Smetters. 2007. "Does Social Security Privatization Produce Efficiency Gains?" *The Quarterly Journal of Economics* 122 (4):1677–1719. 7, 20, xxiv

Rogerson, Richard and Johanna Wallenius. 2013. "Nonconvexities, Retirement, and the Elasticity of Labor Supply." *The American Economic Review* 103 (4):1445–1462. 4, 5, 22, xxi

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1):205–229. 4, 13, viii

———. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *The Quarterly Journal of Economics* 117 (3):1039–1073. 3

Saez, Emmanuel and Stefanie Stantcheva. 2016. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review* 106 (1):24–45. 4, viii

SSA. 2014. "The Annual Statistical Supplement to the Social Security Bulletin." *Social Security Administration Publication* 13 (11700). 21, 30

Waldron, Hilary. 2013. "Mortality Differentials by Lifetime Earnings Decile: Implications for Evaluations of Proposed Social Security Law Changes." *Social Security Bulletin* 73 (1):1–37. 31, 32

Weinzierl, Matthew. 2011. "The Surprising Power of Age-Dependent Taxes." *Review of Economic Studies* 78 (4):1490–1518. 4

Online Appendix - Not For Publication

This online Appendix contains three sections: section A provides proofs omitted in the main text; section B shows the results in the general setup with risk aversion, discounting, heterogeneous life span, and overlapping generations; section C provides further details of the quantitative analysis and robustness checks.

A Proofs for Section 3

We suppress explicit dependence on θ throughout the appendix whenever it does not inhibit clarity.

A.1 Existence of retirement age

We first show that there must exist a retirement age if in the constrained optimum the Lagrange multiplier associated with incentive constraints (8) is non-negative. This can be interpreted as constraints (3) binding only downward, which is guaranteed by social welfare being redistributive from individuals with higher productivity toward those with lower productivity, i.e., $G(\theta) \geq F(\theta)$.

Proposition 5 *Suppose that in the constrained-efficient allocation the Lagrange multiplier on (8) is non-negative, i.e., (3) is binding only for $\hat{\theta} \leq \theta$. Then there exists age $R(\theta)$ such that individuals of type θ prefer to work if and only if $t < R(\theta)$.*

Proof.

Suppose that for an individual of type θ there exists an interval $[t_1, t_2]$ such that for values of t in this interval $y(t,\theta) = 0$ and another interval $[t_2, t_3]$ such that $y(t,\theta) > 0$. We show that any such allocation can be improved upon.

Consider a perturbation of the allocation for which $\tilde{y}(t,\theta) > 0$ for $t \in [t_1, t_1 + \epsilon]$ and $\tilde{y}(t,\theta) = 0$ for $t \in [t_2, t_2 + \epsilon]$. Furthermore, suppose that $\forall t \in [t_1, t_1 + \epsilon]$, $\tilde{y}(t,\theta) = \xi y(t + t_2 - t_1, \theta)$ where ξ is such that

$$\int_{t_1}^{t_1+\epsilon} \left(\frac{\tilde{y}(t,\theta)}{\varphi(t,\theta)}\right)^{1+1/\varepsilon} dt = \int_{t_2}^{t_2+\epsilon} \left(\frac{y(t,\theta)}{\varphi(t,\theta)}\right)^{1+1/\varepsilon} dt$$

Note that this perturbation keeps the length of working life and the disutility of leisure unchanged. However, since any nonworking must occur after $t^*(\theta)$ and as a result $\varphi(t_1,\theta) >$

$\varphi(t_2,\theta)$, the total output from type θ must increase. Therefore, if the allocations remain incentive compatible, such perturbation improves welfare.

In order to show incentive compatibility of the perturbed allocation, it is sufficient to show that

$$\int_{t_1}^{t_1+\epsilon}\left(\frac{\tilde{y}(t,\theta)}{\varphi(t,\theta')}\right)^{1+1/\varepsilon}dt > \int_{t_2}^{t_2+\epsilon}\left(\frac{y(t,\theta)}{\varphi(t,\theta')}\right)^{1+1/\varepsilon}dt, \forall\theta' > \theta \tag{16}$$

This implies that the value of reporting θ for all types $\theta' > \theta$ goes down in response to this perturbation and completes the argument since by assumption only downward incentive constraints bind.

To show (16), note that from Assumption 1 we have that

$$\frac{\varphi(t',\theta)}{\varphi(t,\theta)} < \frac{\varphi(t',\theta')}{\varphi(t,\theta')}, \theta' > \theta, t' > t$$

This inequality implies that

$$\left(\frac{\varphi(t',\theta)}{\varphi(t',\theta')}\right)^{1+1/\varepsilon} < \left(\frac{\varphi(t,\theta)}{\varphi\left(t,\theta'\right)}\right)^{1+1/\varepsilon}, \theta' > \theta, t' > t$$

Hence, multiplying the integrands in (16) by $\left(\frac{\varphi(t,\theta')}{\varphi(t,\theta)}\right)^{1+1/\varepsilon}$, makes the left-hand side larger, i.e., we must have that

$$\int_{t_1}^{t_1+\epsilon}\left(\frac{\tilde{y}(t,\theta)}{\varphi(t,\theta)}\right)^{1+1/\varepsilon}\left(\frac{\varphi(t,\theta)}{\varphi\left(t,\theta'\right)}\right)^{1+1/\varepsilon}dt > \int_{t_2}^{t_2+\epsilon}\left(\frac{y(t,\theta)}{\varphi(t,\theta)}\right)^{1+1/\varepsilon}\left(\frac{\varphi(t,\theta)}{\varphi\left(t,\theta'\right)}\right)^{1+1/\varepsilon}dt$$

This inequality implies (16) and concludes the proof. ■

A.2 Proof of Proposition 1

First note that when $\tilde{\varphi}\left(t,\theta\right)$ is the productivity profile and $\tilde{\eta}\left(\theta\right)$ is the fixed cost of working, optimal hours worked and the retirement decision under public information are given by

$$\psi\frac{y\left(t,\theta\right)^{1/\varepsilon}}{\tilde{\varphi}\left(t,\theta\right)^{1+1/\varepsilon}} = 1 \tag{17}$$

$$y\left(R\left(\theta\right),\theta\right) = \psi\frac{y\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}{\left(1+1/\varepsilon\right)\tilde{\varphi}\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}+\tilde{\eta}\left(\theta\right) \tag{18}$$

The planning problem with private information is to maximize $\int U(\theta)dG(\theta)$ subject to

$$U(\theta)=c(\theta)\bar{T}-\int_0^{R(\theta)}v(l(t,\theta))dt-(R(\theta)-t^*(\theta))\eta(\theta),$$

as well as (8) and (2). Suppressing θ, the first-order conditions are given by

$$g-\alpha-\mu' = 0 \tag{19}$$

$$\alpha-\lambda f = 0 \tag{20}$$

$$\forall t\leq R, -\psi\frac{y\left(t\right)^{1/\varepsilon}}{\varphi\left(t\right)^{1+1/\varepsilon}}\alpha+\lambda f-\mu\psi\left(1+1/\varepsilon\right)\frac{y\left(t\right)^{1/\varepsilon}}{\varphi\left(t\right)^{1+1/\varepsilon}}\frac{\varphi_\theta\left(t\right)}{\varphi\left(t\right)} = 0 \tag{21}$$

$$-\left[\frac{\psi}{\left(1+1/\varepsilon\right)}\frac{y\left(R\right)^{1+1/\varepsilon}}{\varphi\left(R\right)^{1+1/\varepsilon}}+\eta\right]\alpha+y\left(R\right)\lambda f-\left[\psi\frac{\varphi_\theta\left(R\right)}{\varphi\left(R\right)}\frac{y\left(R\right)^{1+1/\varepsilon}}{\varphi\left(R\right)^{1+1/\varepsilon}}-\eta'\right]\mu = 0 \tag{22}$$

$$\mu\left(\underline{\theta}\right)=\mu\left(\bar{\theta}\right) = 0,$$

where $\alpha\left(\theta\right)$ is the multiplier on the $U(\theta)$ constraint, λ is the multiplier on feasibility (2), and $\mu\left(\theta\right)$ is the multiplier on incentive constraints (8). Integrating over equation (19) and using the boundary conditions we have

$$\int_{\underline{\theta}}^{\bar{\theta}}g\left(\theta\right)d\theta-\lambda\int_{\underline{\theta}}^{\bar{\theta}}f\left(\theta\right)d\theta=0,$$

and hence $\lambda=1$ since $G\left(\bar{\theta}\right)=F\left(\bar{\theta}\right)=1$. Moreover, we also have

$$\begin{aligned}\mu\left(\theta\right) &= \int_{\underline{\theta}}^{\theta}\left[g\left(\theta'\right)-f\left(\theta'\right)\right]d\theta' \\ &= G\left(\theta\right)-F\left(\theta\right)\end{aligned}$$

and we can rewrite equation (21) as

$$\psi \frac{y(t)^{1/\varepsilon}}{\varphi(t)^{1+1/\varepsilon}} \left[1+(1+1/\varepsilon)\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\varphi_\theta(t)}{\varphi}\right]=1,$$

while (22) becomes

$$\begin{aligned} y(R) &= \left[\psi\frac{\varphi_\theta(R)}{\varphi(R)}\frac{y(R)^{1+1/\varepsilon}}{\varphi(R)^{1+1/\varepsilon}}-\eta'\right]\frac{G(\theta)-F(\theta)}{f(\theta)}+\left[\frac{\psi}{(1+1/\varepsilon)}\frac{y(R)^{1+1/\varepsilon}}{\varphi(R)^{1+1/\varepsilon}}+\eta\right] \\ &= \psi\frac{y(R)^{1+1/\varepsilon}}{(1+1/\varepsilon)\varphi(R)^{1+1/\varepsilon}}\left[1+(1+1/\varepsilon)\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\varphi_\theta(R)}{\varphi(R)}\right] \\ &\quad +\eta-\eta'\frac{G(\theta)-F(\theta)}{f(\theta)}. \end{aligned}$$

Defining modified productivities and fixed costs

$$\begin{aligned} \tilde{\varphi}(t,\theta) &= \varphi(t,\theta)\left[1+(1+1/\varepsilon)\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)}\right]^{-\frac{\varepsilon}{1+\varepsilon}} \\ \tilde{\eta}(\theta) &= \eta(\theta)\left[1-\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\eta'(\theta)}{\eta(\theta)}\right], \end{aligned}$$

it can be easily seen that (17) and (18) are satisfied for these productivity profiles and fixed costs, establishing the claim. ■

A.3 Proof of Proposition 2

The first-order conditions associated with the mechanism design problem are given by (19)-(22) and the boundary conditions. Note that integrating over equations (19) and (20) and using the boundary conditions $\mu(\underline{\theta})=\mu(\bar{\theta})=0$ implies that $\lambda=1$.

We can then write the third condition (21) as

$$\psi\frac{y(t)^{1/\varepsilon}}{\varphi(t)^{1+1/\varepsilon}}\left[1+(1+1/\varepsilon)\frac{G(\theta)-F(\theta)}{f(\theta)}\frac{\varphi_\theta}{\varphi}\right]=1\rightarrow y(t)=\psi^{-\varepsilon}\tilde{\varphi}(t,\theta)^{1+\varepsilon}$$

Replacing the above in (22), we have

$$\tilde{\varphi}(R(\theta),\theta)^{1+\varepsilon}=\tilde{\eta}(\theta)(1+\varepsilon)\psi^{\varepsilon}$$

This proves the claim. ■

A.4 Proof of Proposition 3

It is sufficient to show equation (11) in the proposition since the rest of the argument is given in the main text. Note that the definitions of wedges (5) and (6) imply

$$\begin{aligned}\tau_R(\theta)\, y(R(\theta),\theta) &= y(R(\theta),\theta) - \left[\frac{\psi}{1+1/\varepsilon}\frac{y(R(\theta),\theta)^{1+1/\varepsilon}}{\varphi(R(\theta),\theta)^{1+1/\varepsilon}} + \eta(\theta)\right] \\ \tau_y(R(\theta),\theta)\, y(R(\theta),\theta) &= y(R(\theta),\theta) - \psi\frac{y(R(\theta),\theta)^{1+1/\varepsilon}}{\varphi(R(\theta),\theta)^{1+1/\varepsilon}},\end{aligned}$$

and hence, using (21) and (22) we arrive at

$$\begin{aligned}&\tau_R(\theta)\, y(R(\theta),\theta) - \frac{1}{1+1/\varepsilon}\tau_y(R(\theta),\theta)\, y(R(\theta),\theta) \\ &= (1+\varepsilon)\, y(R(\theta),\theta) - \eta(\theta) \\ &= -\frac{G(\theta)-F(\theta)}{f(\theta)}\eta'(\theta),\end{aligned}$$

where the last equality follows from (17) and (18). ■

A.5 Proof of Lemma 1

Since $y(t,\theta)$ is increasing and continuous in θ, then the range of $y(t,\cdot):\Theta\mapsto\mathbb{R}$ is an interval, $\left[\underline{y}(t),\bar{y}(t)\right]$. The envelope condition associated with the maximization in (13) is given by

$$v'(\theta) = \psi\frac{\varphi_\theta(t,\theta)}{\varphi(t,\theta)}\frac{y(t,\theta)^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}},$$

where $v(\theta) = y(t,\theta) - \mathcal{T}(t,y(t,\theta)) - \frac{\psi}{1+1/\varepsilon}\frac{y(t,\hat{\theta})^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}}$. Hence, we must have

$$v(\theta) = \underline{v} + \int_{\underline{\theta}(t)}^{\theta}\psi\frac{\varphi_\theta(t,\hat{\theta})}{\varphi(t,\hat{\theta})}\frac{y(t,\hat{\theta})^{1+1/\varepsilon}}{\varphi(t,\hat{\theta})^{1+1/\varepsilon}}d\hat{\theta},$$

where $\underline{\theta}(t)$ is the lowest type that works at age t. As a result

$$y(t,\theta) - \mathcal{T}(t,y(t,\theta)) - \frac{\psi}{1+1/\varepsilon}\frac{y(t,\hat{\theta})^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}} = \underline{v} + \int_{\underline{\theta}(t)}^{\theta}\psi\frac{\varphi_\theta(t,\hat{\theta})}{\varphi(t,\hat{\theta})}\frac{y(t,\hat{\theta})^{1+1/\varepsilon}}{\varphi(t,\hat{\theta})^{1+1/\varepsilon}}d\hat{\theta}$$

This equation uniquely defines $\mathcal{T}(t,\cdot)$ up to a constant. Furthermore, since $y(t,\theta)$ is increasing in θ, standard mechanism design arguments establish that for any such $\mathcal{T}$, equation (13) is indeed satisfied (see, e.g., Fudenberg and Tirole (1991)). ∎

A.6 Proof of Lemma 2

Given the tax schedule and benefit formula, a household of type θ's optimization problem is given by

$$\begin{aligned} \max_{R(\theta),y(t)} \quad & \int_0^{R(\theta)} \left[y(t,\theta) - \mathcal{T}(t,y(t,\theta))\right] dt + b(R) \\ & - \int_0^{R(\theta)} \left[\frac{\psi}{1+1/\varepsilon} \frac{y(t,\theta)^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}}\right] dt - \eta(\theta)\left(R(\theta) - t^*(\theta)\right) \end{aligned} \tag{23}$$

We refer to the solution for (23) as $\tilde{y}(t,\theta)$ and $\tilde{R}(\theta)$. Our goal is to show that $\tilde{y}(t,\theta) = y(t,\theta)$ and $\tilde{R}(\theta) = R(\theta)$. We start by showing the following lemma:

Lemma 3 *If $\tilde{y}(t,\theta) \geq 0$, then $\tilde{y}(t,\theta) = y(t,\theta)$.*

Proof. The proof simply follows from the definition of $\mathcal{T}$. Since $\theta = \arg\max_{\hat{\theta}} y\left(t,\hat{\theta}\right) - \mathcal{T}\left(t, y\left(t,\hat{\theta}\right)\right) - v\left(\frac{y\left(t,\hat{\theta}\right)}{\varphi(t,\theta)}\right)$, it is optimal for an individual of type θ to choose $y(t,\theta)$ at age t. ∎

Using the above lemma, we can show the following:

Lemma 4 *Choosing $\{y(t,\theta)\}_{t\leq R(\theta)}, R(\theta)$ for an agent of type θ is a local optimum for an individual of type θ in (23).*

Proof. Suppose on the contrary that the individual chooses $R\left(\hat{\theta}\right) \neq R(\theta)$, then given the definition of b, the utility for the household is given by

$$\begin{aligned} & \int_0^{R(\hat{\theta})} \left[y(t,\theta) - \mathcal{T}(t,y(t,\theta))\right] dt - \int_0^{R(\hat{\theta})} \left[\frac{\psi}{1+1/\varepsilon} \frac{y(t,\theta)^{1+1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}}\right] dt - \eta(\theta)\left(R(\theta) - t^*(\theta)\right) \\ & \quad + c\left(\hat{\theta}\right)\bar{T} - \int_0^{R(\hat{\theta})} \left[y\left(t,\hat{\theta}\right) - \mathcal{T}\left(t, y\left(t,\hat{\theta}\right)\right)\right] dt \end{aligned} \tag{24}$$

Taking a derivative with respect to $\hat{\theta}$, we have

$$\left[y\left(R\left(\hat{\theta}\right),\theta\right)-\mathcal{T}\left(R\left(\hat{\theta}\right),y\left(R\left(\hat{\theta}\right),\theta\right)\right)-\frac{\psi}{1+1/\varepsilon}\frac{y\left(R\left(\hat{\theta}\right),\theta\right)^{1+1/\varepsilon}}{\varphi\left(R\left(\hat{\theta}\right),\theta\right)^{1+1/\varepsilon}}-\eta\left(\theta\right)\right]R'\left(\hat{\theta}\right)$$
$$+c'\left(\hat{\theta}\right)\bar{T}-\left[y\left(R\left(\hat{\theta}\right),\hat{\theta}\right)-\mathcal{T}\left(R\left(\hat{\theta}\right),y\left(R\left(\hat{\theta}\right),\hat{\theta}\right)\right)\right]R'\left(\hat{\theta}\right)$$
$$-\int_0^{R\left(\hat{\theta}\right)}\frac{\partial}{\partial\theta}y\left(t,\hat{\theta}\right)\left[1-\frac{\partial}{\partial y}\mathcal{T}\left(t,y\left(t,\hat{\theta}\right)\right)\right]dt$$

Evaluating the above expression when $\hat{\theta}=\theta$,

$$c'\left(\theta\right)\bar{T}-\left[\frac{\psi}{1+1/\varepsilon}\frac{y\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}{\varphi\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}+\eta\left(\theta\right)\right]R'\left(\theta\right)-\int_0^{R(\theta)}\frac{\partial}{\partial\theta}y\left(t,\theta\right)\left[1-\frac{\partial}{\partial y}\mathcal{T}\left(t,y\left(t,\theta\right)\right)\right]dt,$$

and by static incentive compatibility (13) the above expression becomes

$$c'\left(\theta\right)\bar{T}-\left[\frac{\psi}{1+1/\varepsilon}\frac{y\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}{\varphi\left(R\left(\theta\right),\theta\right)^{1+1/\varepsilon}}+\eta\left(\theta\right)\right]R'\left(\theta\right)-\int_0^{R(\theta)}\psi\frac{y\left(t,\theta\right)^{1/\varepsilon}}{\varphi\left(t,\theta\right)^{1+1/\varepsilon}}\frac{\partial}{\partial\theta}y\left(t,\theta\right)dt,$$

which is zero by incentive compatibility of the original allocation. This implies that $\hat{\theta}=\theta$ is a local optimum point of the function (24), and hence $R=R\left(\hat{\theta}\right)$ is a local optimum of (23). ■

Together the above lemmas establish the claim in Lemma 2. ■

B Proofs for the general setup of Section 2

We extend here the qualitative results in the main text to the general setup with risk aversion and discounting.

B.1 Equivalence

We start by showing that the equivalence result generalizes in an intuitive way and then show that the sufficient-statistic-type expressions remain unchanged.

The first-order conditions (19)-(22) associated with the mechanism design problem now become

$$g-\alpha-\mu'=0 \tag{25}$$

$$\alpha u'(c)-\lambda f=0 \tag{26}$$

$$\forall t \leq R, \quad -\psi \frac{y(t)^{1/\varepsilon}}{\varphi(t)^{1+1/\varepsilon}}\alpha + \lambda f - \mu\psi\left(1+1/\varepsilon\right)\frac{y(t)^{1/\varepsilon}}{\varphi(t)^{1+1/\varepsilon}}\frac{\varphi_\theta(t)}{\varphi(t)} = 0 \tag{27}$$

$$-\left[\frac{\psi}{1+1/\varepsilon}\frac{y(R)^{1+1/\varepsilon}}{\varphi(R)^{1+1/\varepsilon}} + \eta\right]\alpha + y(R)\lambda f - \left[\psi\frac{\varphi_\theta(R)}{\varphi(R)}\frac{y(R)^{1+1/\varepsilon}}{\varphi(R)^{1+1/\varepsilon}} - \eta'\right]\mu = 0 \tag{28}$$

$$\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$$

Note that the previous arguments for treating consumption levels as constant over time are unaffected. Thus the first two equations (25) and (26) combined with the boundary conditions give

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}\left[g(\theta') - \frac{\lambda f(\theta')}{u'(c(\theta'))}\right]d\theta'$$

Recall that λ is the marginal value of public funds for the planner. Following Saez (2001) and Saez and Stantcheva (2016), let us denote $\bar{g}(\theta)$ the ratio of the marginal value of public funds for the planner to utilitarian marginal welfare from per-capita consumption of θ-type individuals, i.e., $\bar{g}(\theta) = \frac{\lambda f(\theta)}{u'(c(\theta))}$. In other words, a utilitarian planner is indifferent between $\bar{g}(\theta)$ more public funds and a marginal decrease in θ-type consumption. The larger $\bar{g}(\theta)$, the less a utilitarian planner values consumption by θ-types and therefore $\bar{g}(\theta)$ is a parameter reflecting exogenous motive to redistribute from θ. The same motive for all types up to θ is then $\bar{G}(\theta) = \int_{\underline{\theta}}^{\theta}\frac{\lambda f(\theta')}{u'(c(\theta'))}d\theta'$. Thus

$$\mu(\theta) = G(\theta) - \bar{G}(\theta).$$

We can then re-write equation (27) as

$$\frac{\psi y(t)^{1/\varepsilon}}{\varphi(t)^{1+1/\varepsilon}}\left[1 + \left(1+\frac{1}{\varepsilon}\right)\frac{G(\theta)-\bar{G}(\theta)}{\bar{g}(\theta)}\frac{\varphi_\theta(t)}{\varphi}\right]\frac{\bar{g}(\theta)}{\lambda f(\theta)} = 1,$$

and (28) becomes

$$\begin{aligned} y(R) \quad = \quad & \frac{\psi y(R)^{1+1/\varepsilon}}{(1+1/\varepsilon)\varphi(R)^{1+1/\varepsilon}}\left[1 + \left(1+\frac{1}{\varepsilon}\right)\frac{G(\theta)-\bar{G}(\theta)}{\bar{g}(\theta)}\frac{\varphi_\theta(R)}{\varphi(R)}\right]\frac{\bar{g}(\theta)}{\lambda f(\theta)} \\ & + \left[\eta - \eta'\frac{G(\theta)-\bar{G}(\theta)}{\bar{g}(\theta)}\right]\frac{\bar{g}(\theta)}{\lambda f(\theta)}. \end{aligned}$$

Then the equivalence result in the general environment follows if the virtual productivities and fixed costs are given by

$$\begin{aligned}
\tilde{\varphi}\left(t,\theta\right) &= \varphi(t,\theta)\left[\left(1+\left(1+\frac{1}{\varepsilon}\right)\frac{G\left(\theta\right)-\bar{G}\left(\theta\right)}{\bar{g}\left(\theta\right)}\frac{\varphi_{\theta}\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)\frac{\bar{g}(\theta)}{\lambda f(\theta)}\right]^{-\frac{\varepsilon}{1+\varepsilon}}, \\
\tilde{\eta}\left(\theta\right) &= \eta\left(\theta\right)\left[\left(1-\frac{G\left(\theta\right)-\bar{G}\left(\theta\right)}{\bar{g}\left(\theta\right)}\frac{\eta'\left(\theta\right)}{\eta\left(\theta\right)}\right)\frac{\bar{g}(\theta)}{\lambda f(\theta)}\right].
\end{aligned}$$

That is, to maintain the full information optimality (17) and (18) the modification of productivities and fixed costs must be given by the above.

Intuitively, these are analogous modifications to the ones we derived in the main text for the quasi-linear example with the exception of what is being used as a benchmark of utilitarian redistribution. With risk aversion it is no longer enough to simply compare redistributive motives, G, against the distribution of types, F, as was the case in the quasi-linear example. Since risk aversion introduces curvature into the utility of consumption, even a utilitarian benchmark would take that into account trading it off against funds available to the rest of the population. This is precisely what $\bar{G}$ captures.

On the other hand, if curvature is removed the planner is able to transform public funds one for one into the utility of consumption of a given type, making $\bar{G}$ equal to F and making the virtual types above identical to the ones derived in the main text.

Finally, using these definitions of virtual types we can re-write (28) as

$$\tilde{\varphi}(R(\theta),\theta)^{1+\varepsilon} = \tilde{\eta}(\theta)\left(1+\varepsilon\right)\psi^{\varepsilon}$$

Thus, the sufficient statistic expression is unchanged and as in the main text the constrained-efficient retirement age is increasing if and only if $\frac{\partial}{\partial\theta}\frac{\tilde{\varphi}(t,\theta)^{1+\varepsilon}}{\tilde{\eta}(\theta)}\Big|_{t=R(\theta)} \geq 0$.

B.2 Retirement incentives vs. hours incentives

Note first that an implication of the first-order conditions (25) and (26) combined with the boundary conditions can be re-written as

$$\mu\left(\theta\right) = \int_{\theta}^{\bar{\theta}}\left[\frac{\lambda f\left(\theta'\right)}{u'\left(c\left(\theta'\right)\right)} - g\left(\theta'\right)\right]d\theta'$$

Multiplying the third equation (27) for $t = R\left(\theta\right)$ by $\frac{y(R)}{1+1/\varepsilon}$ and subtracting from (28) we have

$$-\eta\alpha + \left(1+\varepsilon\right)y\left(R\right)\lambda f + \eta'\mu = 0$$

Hence,

$$\begin{aligned}(1+\varepsilon)\, y(R) &= \frac{\eta}{u'(c(\theta))} - \eta' \frac{\mu}{\lambda f} \\ &= \frac{\eta}{u'(c(\theta))} - \eta' \frac{1-F(\theta)}{f(\theta)} \int_{\theta}^{\bar{\theta}} \left[\frac{1}{u'(c(\theta'))} - \frac{g(\theta')}{\lambda f(\theta')} \right] \frac{dF(\theta')}{1-F(\theta)} \end{aligned} \tag{29}$$

Note that the wedges are given by

$$\begin{aligned}\tau_R(\theta)\, y(R(\theta),\theta) &= y(R(\theta),\theta) - \frac{1}{u'(c(\theta))} \left[\frac{\psi}{1+1/\varepsilon} \frac{y(R(\theta),\theta)^{1+1/\varepsilon}}{\varphi(R(\theta),\theta)^{1+1/\varepsilon}} + \eta(\theta) \right] \\ \tau_y(R(\theta),\theta)\, y(R(\theta),\theta) &= y(R(\theta),\theta) - \frac{1}{u'(c(\theta))} \psi \frac{y(R(\theta),\theta)^{1+1/\varepsilon}}{\varphi(R(\theta),\theta)^{1+1/\varepsilon}} \end{aligned}$$

and hence

$$\begin{aligned} &\tau_R(\theta)\, y(R(\theta),\theta) - \frac{1}{1+1/\varepsilon} \tau_y(R(\theta),\theta)\, y(R(\theta),\theta) \\ &= (1+\varepsilon)\, y(R(\theta),\theta) - \frac{\eta(\theta)}{u'(c(\theta))} \\ &= -\eta' \frac{1-F(\theta)}{f(\theta)} \int_{\theta}^{\bar{\theta}} \left[\frac{1}{u'(c(\theta'))} - \frac{g(\theta')}{\lambda f(\theta')} \right] \frac{dF(\theta')}{1-F(\theta)}, \end{aligned}$$

where the last equality follows from (29). ■

B.3 Implementation

Here we construct tax and pension benefit policies to implement an incentive-compatible allocation in our general model. Our goal is to construct an age-dependent tax function, $\mathcal{T}(t,y)$, as well as a benefit function given by present value $b(R,Y)$, where Y is a measure of life-time earnings. Throughout this section we keep the assumption that $v(l) = \psi l^{1+1/\varepsilon}/(1+1/\varepsilon)$. We make the following assumptions about the implemented allocation:

Assumption 2 *The function $y(t,\cdot) : \Theta \longmapsto \mathbb{R}$ is increasing in θ.*

This assumption implies that a tax function that only depends on income at age t is well defined. Note that in this implementation each individual is solving the following optimization problem:

$$\max_{y(t),R,c} \int_0^{\bar{T}} e^{-\rho t} u(c)\, dt - \int_0^R e^{-\rho t} \left[v\left(y(t) / \varphi(t,\theta) \right) + \eta(t,\theta) \right] dt$$

subject to

$$\int_0^{\bar{T}} e^{-\rho t} c dt = \int_0^R e^{-\rho t} \left[y(t) - \mathcal{T}(t, y(t)) \right] dt + e^{-\rho R} b\left(R, \frac{1}{R} \int_0^R y(t)\, dt \right)$$

As in the main text, we start by constructing the tax function. For any tax function, $\mathcal{T}(t,y)$, we consider the following auxiliary optimization problem:

$$\mathcal{V}(b,R,Y;\theta,\mathcal{T}) = \max_{\{\tilde{y}(t)\}_{t\in[0,R]},c} \int_0^{\bar{T}} e^{-\rho t} u(c)\, dt - \int_0^R e^{-\rho t} \left[v\left(\tilde{y}(t) / \varphi(t,\theta) \right) + \eta(t,\theta) \right] dt \tag{30}$$

subject to

$$\begin{aligned} \int_0^{\bar{T}} e^{-\rho t} c dt &= \int_0^{\bar{T}} e^{-\rho t} \left[\tilde{y}(t) - \mathcal{T}(t, \tilde{y}(t)) \right] dt + e^{-\rho R} b \\ Y &= \frac{1}{R} \int_0^{\bar{T}} \tilde{y}(t)\, dt \end{aligned} \tag{31}$$

We say a tax function, $\mathcal{T}(t,y)$, *partially implements* the allocation $c(\theta), \{y(t,\theta)\}_{t\in[0,R(\theta)]}, R(\theta)$, if $\{y(t,\theta)\}_{t\in[0,R(\theta)]}$ is the solution to (30) when $R = R(\theta)$,

$$b = e^{\rho R(\theta)} \left[\int_0^{\bar{T}} e^{-\rho t} c(\theta)\, dt - \int_0^{R(\theta)} e^{-\rho t} \left[y(t,\theta) - \mathcal{T}(t, y(t,\theta)) \right] dt \right],$$

and $Y = Y(\theta) = \frac{1}{R(\theta)} \int_0^{R(\theta)} y(t,\theta)\, dt$. The following lemma establishes that such a tax function exists.

Lemma 5 *Consider the constrained-efficient allocation with extended earnings* $\left\{ c(\theta), \{y(t,\theta)\}_{t\in[0,1]}, R(\theta) \right\}_{\theta\in\Theta}$. *Then*

1. *There exists a tax function* $\mathcal{T}(t,y)$ *so that* $\{y(t,\theta)\}_{t\in[0,R(\theta)]}$ *is a local optimum in (30) for* $R = R(\theta)$, $b = e^{\rho R(\theta)} \left[\int_0^{\bar{T}} e^{-\rho t} c(\theta)\, dt - \int_0^{R(\theta)} e^{-\rho t} \left[y(t,\theta) - \mathcal{T}(t, y(t,\theta)) \right] dt \right]$, *and* $Y = Y(\theta) = \frac{1}{R(\theta)} \int_0^{R(\theta)} y(t,\theta)\, dt$.

2. *Suppose that for all* t, $\tau_y(t,\theta) < 1$ *and that* $\frac{y(t,\theta)^{1+\frac{1}{\varepsilon}}}{1-\tau_y(t,\theta)}$ *and* $y(t,\theta)$ *move together. That is,*

$y(t,\theta) > y(t,\theta')$ if and only if $\frac{y(t,\theta)^{1+\frac{1}{\varepsilon}}}{1-\tau_y(t,\theta)} > \frac{y(t,\theta')^{1+\frac{1}{\varepsilon}}}{1-\tau_y(t,\theta')}$. Then there exists a tax function that partially implements the allocation $c(\theta), \{y(t,\theta)\}_{t\in[0,R(\theta)]}, R(\theta)$.

Proof. 1. We prove the first part by constructing one such tax function. Let $\tau_y(t,\theta)$ be the labor wedge faced by individual θ at age t. We define the tax function $\mathcal{T}$ as follows:

$$\begin{aligned} \mathcal{T}(t,y) &= \underline{\mathcal{T}}(t) + \int_{\underline{y}(t)}^{\min\{y,\bar{y}(t)\}} \tau_y\left(t,\hat{\theta}(y,t)\right) dy \\ &\quad + (y - \min\{y,\bar{y}(t)\}), \forall y \in \left[\underline{y}(t), \infty\right) \\ \mathcal{T}(t,y) &= \underline{\mathcal{T}}(t), \forall y \in \left[0, \underline{y}(t)\right), \end{aligned} \tag{32}$$

where $\underline{y}(t) = \min_{\theta\in\Theta} y(t,\theta)$, $\bar{y}(t) = \max_{\theta\in\Theta} y(t,\theta)$, and $y\left(t,\hat{\theta}(t,y)\right) = y$. Note that the function $\hat{\theta}(t,y)$ is well-defined since $y(t,\theta)$ is a one-to-one function of θ. In addition, $\underline{\mathcal{T}}(t)$ is an arbitrary function of age. The intercept of the tax function $\underline{\mathcal{T}}(t)$ is an arbitrary continuous function. Note that (32) implies that the marginal tax rate is 1 for values of y above $\bar{y}(t)$. Given our construction, it is clear that $\{y(t,\theta)\}_{t\in[0,R(\theta)]}$ satisfies the first-order conditions associated with (30) when the last constraint is slack – since marginal tax rate evaluated at $y(t,\theta)$ is equal to the labor wedge.

2. To prove the second claim, we conjecture that the constraint (31) is slack. Then any solution to (30) must satisfy the following local optimality condition:

$$u'(c)\left(1 - \mathcal{T}_y(t,\tilde{y}(t))\right) = \psi \frac{\tilde{y}(t)^{1/\varepsilon}}{\varphi(t,\theta)^{1+1/\varepsilon}}$$

We can rewrite this equation as

$$\left(1 - \mathcal{T}_y(t,\tilde{y}(t))\right)\tilde{y}(t)^{-1/\varepsilon} = \psi \frac{u'(c)}{\varphi(t,\theta)^{1+1/\varepsilon}}$$

By the assumption stated above, the left-hand side of the above equation is decreasing in $\tilde{y}(t)$. Now suppose that $c > c(\theta)$, then $u'(c) < u'(c(\theta))$ and hence $\tilde{y}(t) < y(t,\theta)$. Since marginal tax rates are less than 1,

$$\int_0^{R(\theta)} e^{-\rho t}\left[\tilde{y}(t) - \mathcal{T}(t,\tilde{y}(t))\right] dt \le \int_0^{R(\theta)} e^{-\rho t}\left[y(t,\theta) - \mathcal{T}(t,y(t,\theta))\right] dt,$$

which implies that the budget constraint cannot hold. A similar argument implies that $c < c(\theta)$ cannot hold, and hence the unique solution to a relaxed version of (30) is

$\{y(t,\theta)\}_{t\in[0,R(\theta)]}, c(\theta)$ which satisfies (31). This establishes the claim. ■

The properties assumed above are satisfied in all of our numerical simulations. Note also that there are potentially many tax functions that partially implement the constrained-efficient allocations. Our construction of benefits below works for all tax functions that partially implement the constrained-efficient allocations.

Consider a tax function $\mathcal{T}(t,y)$ that partially implements the constrained-efficient allocations and its associated pension benefits defined by

$$\hat{b}(\theta) = e^{\rho R(\theta)} \left[\int_0^{\bar{T}} e^{-\rho t} c(\theta)\, dt - \int_0^{R(\theta)} e^{-\rho t} \left[y(t,\theta) - \mathcal{T}(t, y(t,\theta)) \right] dt \right]$$

Our goal is to find a function $b(R,Y)$ such that $b(R(\theta), Y(\theta)) = \hat{b}(\theta)$ and

$$\mathcal{V}(b(R,Y), R, Y; \theta, \mathcal{T}) \leq \mathcal{V}(b(\theta), R(\theta), Y(\theta); \theta, \mathcal{T})$$

This would imply that an individual of type θ finds it optimal to choose $R(\theta)$ and $Y(\theta)$. Lemma 5 then implies that having chosen $R(\theta)$ and $Y(\theta)$, the optimal choice for earnings and consumption is the constrained-efficient allocation. The following proposition establishes that function $b(\cdot,\cdot)$ exists.

Proposition 6 *Suppose that Θ is a compact set. There exists a function $b(\cdot,\cdot)$ such that $b(R(\theta), Y(\theta)) = \hat{b}(\theta)$ and $\mathcal{V}(b(R,Y), R, Y; \theta, \mathcal{T}) \leq \mathcal{V}(b(\theta), R(\theta), Y(\theta); \theta, \mathcal{T})$.*

Proof. Our proof of this proposition is in two steps. First, we show that

$$\theta \in \arg\max_{\theta'} \mathcal{V}\left(b\left(\theta'\right), R\left(\theta'\right), Y\left(\theta'\right); \theta, \mathcal{T}\right) \tag{33}$$

We show this by first showing that $\theta' = \theta$ is the local optimum in the above optimization. To see this, we take a derivative of the above function and apply envelope condition to get

$$\begin{aligned} \frac{\partial}{\partial \theta'} \mathcal{V}\left(b\left(\theta'\right), R\left(\theta'\right), Y\left(\theta'\right); \theta, \mathcal{T}\right)\Big|_{\theta'=\theta} &= \\ \mathcal{V}_b b'(\theta) + \mathcal{V}_R R'(\theta) + \mathcal{V}_Y Y'(\theta) &= \end{aligned}$$

$$
\begin{aligned}
&e^{-\rho R(\theta)}u'\left(c\left(\theta\right)\right)b'\left(\theta\right)\\
&+\left\{-e^{-\rho R(\theta)}\left[v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)+\eta\left(\theta\right)\right]\right.\\
&+u'\left(c\left(\theta\right)\right)\left[e^{-\rho R(\theta)}\left(y\left(R\left(\theta\right),\theta\right)-\mathcal{T}\left(R\left(\theta\right),y\left(R\left(\theta\right),\theta\right)\right)\right)-\rho e^{-\rho R(\theta)}b\left(\theta\right)\right]\\
&\left.+\mu\left[\frac{1}{R\left(\theta\right)^2}\int_0^{R(\theta)}y\left(t,\theta\right)dt-\frac{1}{R\left(\theta\right)}y\left(R\left(\theta\right),\theta\right)\right]\right\}R'\left(\theta\right)-\mu Y'\left(\theta\right)=
\end{aligned}
$$

$$
\begin{aligned}
&e^{-\rho R(\theta)}u'\left(c\left(\theta\right)\right)\begin{bmatrix}\rho R'\left(\theta\right)b\left(\theta\right)+e^{\rho R(\theta)}\int_0^{\bar{T}}e^{-\rho t}c'\left(\theta\right)dt\\ -e^{\rho R(\theta)}\int_0^{R(\theta)}e^{-\rho t}\left[1-\mathcal{T}_y\left(t,y\left(t,\theta\right)\right)\right]y_\theta\left(t,\theta\right)dt\\ -\left[y\left(R\left(\theta\right),\theta\right)-\mathcal{T}\left(R\left(\theta\right),y\left(R\left(\theta\right),\theta\right)\right)\right]R'\left(\theta\right)\end{bmatrix}\\
&+\left\{-e^{-\rho R(\theta)}\left[v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)+\eta\left(\theta\right)\right]\right.\\
&\quad+u'\left(c\left(\theta\right)\right)\left[e^{-\rho R(\theta)}\left(y\left(R\left(\theta\right),\theta\right)-\mathcal{T}\left(R\left(\theta\right),y\left(R\left(\theta\right),\theta\right)\right)\right)-\rho e^{-\rho R(\theta)}b\left(\theta\right)\right]\\
&\quad\left.+\mu\left[\frac{1}{R\left(\theta\right)^2}\int_0^{R(\theta)}y\left(t,\theta\right)dt-\frac{1}{R\left(\theta\right)}y\left(R\left(\theta\right),\theta\right)\right]\right\}R'\left(\theta\right)\\
&-\mu\left\{\left[\frac{1}{R\left(\theta\right)^2}\int_0^{R(\theta)}y\left(t,\theta\right)dt-\frac{1}{R\left(\theta\right)}y\left(R\left(\theta\right),\theta\right)\right]R'\left(\theta\right)+\frac{1}{R\left(\theta\right)}\int_0^{R(\theta)}y_\theta\left(t,\theta\right)dt\right\}=
\end{aligned}
$$

$$
\begin{aligned}
&\int_0^{\bar{T}}e^{-\rho t}u'\left(c\left(\theta\right)\right)c'\left(\theta\right)dt-e^{-\rho R(\theta)}\left[v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)+\eta\left(\theta\right)\right]R'\left(\theta\right)\\
&-\int_0^{R(\theta)}y_\theta\left(t,\theta\right)\left[e^{-\rho t}u'\left(c\left(\theta\right)\right)\left(1-\mathcal{T}_y\left(t,y\left(t,\theta\right)\right)\right)+\mu\frac{1}{R\left(\theta\right)}\right]dt=
\end{aligned}
$$

$$
\begin{aligned}
&\int_0^{\bar{T}}e^{-\rho t}u'\left(c\left(\theta\right)\right)c'\left(\theta\right)dt-e^{-\rho R(\theta)}\left[v\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)+\eta\left(\theta\right)\right]R'\left(\theta\right)\\
&-\int_0^{R(\theta)}e^{-\rho t}\frac{y_\theta\left(t,\theta\right)}{\varphi\left(t,\theta\right)}v'\left(\frac{y\left(t,\theta\right)}{\varphi\left(t,\theta\right)}\right)dt,
\end{aligned}
$$

where μ is the Lagrange multiplier on (31) and the last equality follows from the first-order conditions in (30). Note that the final expression above is the same as the local incentive compatibility of the constrained-efficient allocation and hence zero. Therefore, $\theta' = \theta$ is the local optimum of the optimization problem (33). Extensive tedious algebra available upon request shows that under certain conditions $\theta' = \theta$ is also a maximum in (33).

Now, we can construct the function $b(\cdot,\cdot)$. For each value of (R,Y), let $\tilde{b}(\theta)$ be defined by $\mathcal{V}\left(\tilde{b}(\theta), R, Y; \theta, \mathcal{T}\right) = \mathcal{V}(b(\theta), R(\theta), Y(\theta); \theta, \mathcal{T})$. Then, we let $b(R,Y) = \min_\theta \tilde{b}(\theta)$. Since $\mathcal{V}(b, R, Y; \theta, \mathcal{T})$ is increasing in b, we must have

$$\mathcal{V}(b(R,Y), R, Y; \theta, \mathcal{T}) \leq \mathcal{V}\left(\tilde{b}(\theta), R, Y; \theta, \mathcal{T}\right) = \mathcal{V}(b(\theta), R(\theta), Y(\theta); \theta, \mathcal{T})$$

Note that by Assumption 2, since $Y(\theta)$ is strictly increasing in θ, $b(\cdot,\cdot)$ is well-defined. This completes the proof. ■

B.4 Heterogeneous life span

The formal statement of the mechanism design problem discussed in Section 5.4 is given by the following problem:

$$\max \int_\Theta \left[\int_0^{\bar{T}(\theta)} e^{-\rho t} u(c(t,\theta))\, dt - \int_0^{R(\theta)} e^{-\rho t} \left[v(y(t,\theta)/\varphi(t,\theta)) + \eta(t,\theta) \right] dt \right] dG(\theta)$$

subject to feasibility

$$\int_\Theta \int_0^{\bar{T}(\theta)} e^{-\rho t} c(t,\theta)\, dt dF(\theta) + H \leq \int_\Theta \int_0^{R(\theta)} e^{-\rho t} y(t,\theta)\, dt dF(\theta)$$

and incentive compatibility

$$\int_0^{\bar{T}(\theta)} e^{-\rho t} u(c(t,\theta))\, dt - \int_0^{R(\theta)} e^{-\rho t} \left[v(y(t,\theta)/\varphi(t,\theta)) + \eta(t,\theta) \right] dt \geq$$

$$\int_0^{\bar{T}(\theta')} e^{-\rho t} u\left(c\left(t,\theta'\right)\right) dt - \int_0^{R(\theta')} e^{-\rho t} \left[v\left(y\left(t,\theta'\right)/\varphi(t,\theta)\right) + \eta(t,\theta) \right] dt, \forall \theta, \theta'$$

Closely following our baseline analysis we use the first-order approach to incentive compatibility and then numerically verify global incentive compatibility ex post.

B.5 Overlapping generations

We show here that the problem considered in Section 2 is equivalent to the steady state associated with a planning problem of an overlapping generations economy.

Time is continuous, $t \in \mathbb{R}_+ \cup \{0\}$, and at each point in time a generation with a unit mass is born. The individuals born at t live until $t + \overline{T}$. The heterogeneity within a generation is represented by θ, with distribution $F(\theta)$. As in the main text, the production function is given by

$$\mathcal{F}(K(t), L(t)) = rK(t) + L(t),$$

where $\mathcal{F}$ is net output (GDP net of depreciation) and r is net capital income (capital income net of depreciation). The preferences over sequences of consumption, labor supply, and retirement are given by

$$\int_0^{\overline{T}} e^{-\rho a} u(c(a)) da - \int_0^{R(\theta)} e^{-\rho a} \left[v(l(a)) + \eta(a, \theta) \right] da$$

Allocations are given by

- individual allocations $c(t, a, \theta)$, $l(t, a, \theta)$, $R(\theta)$ for all $t \geq -\overline{T}$ and $a \in [\max\{t, 0\}, \overline{T}]$, where t is the time of birth and a is the age of the individual (this includes the initial generations that were born prior to $t = 0$) and
- aggregate capital given by $K(t)$ and aggregate consumption and aggregate effective labor given by

$$\begin{aligned} C(t) &= \int_0^{\overline{T}} \int_{\Theta} c(t-a, a, \theta) dF(\theta) da \\ L(t) &= \int_0^{\overline{T}} \int_{\Theta} \varphi(a, \theta) l(t-a, a, \theta) \mathbf{1}[a \leq R(\theta)] dF(\theta) da \end{aligned}$$

Feasibility requires that

$$C(t) + \hat{H} + \dot{K}(t) = rK(t) + L(t),$$

where $\hat{H}$ is the value of output purchased by the government in each period and is constant over time. Assuming as in the main text that $r = \rho$, the above can be written in its present value form eliminating capital:

$$\int_0^{\infty} e^{-rt} C(t) dt + \frac{\hat{H}}{r} = rK(0) + \int_0^{\infty} e^{-rt} L(t) dt.$$

Disaggregate this into individual consumptions and use integration by parts to arrive at the following:

$$\int_{-\overline{T}}^{\infty} e^{-rt} \int_{\Theta} \int_{\max\{-t,0\}}^{\overline{T}} e^{-ra} c(t, a, \theta) da dF(\theta) dt + \frac{\hat{H}}{r} =$$
$$rK(0) + \int_{-\overline{T}}^{\infty} e^{-rt} \int_{\Theta} \int_{\max\{-t,0\}}^{\overline{T}} e^{-ra} \varphi(t, a, \theta) l(t, a, \theta) da dF(\theta) dt$$

Incentive compatibility is as it is defined in the main text.

The objective function for the planning problem is given by

$$\int_{-\overline{T}}^{\infty} e^{-\lambda t} \int_{\Theta} U(t,\theta) dG(\theta) dt, \tag{34}$$

where $U(t,\theta)$ is the utility of an individual born at t. When $t \geq 0$, this is given by

$$U(t,\theta) = \int_0^{\overline{T}} e^{-\rho t} u(c(t,a,\theta)) da - \int_0^{R(t,\theta)} e^{-\rho a} [v(l(t,a,\theta)) + \eta(a,\theta)] da,$$

while for individuals born at $t \leq 0$, this is given by

$$U(t,\theta) = \int_{-t}^{\overline{T}} e^{-\rho t} u(c(t,a,\theta)) da - \int_{-t}^{R(\theta)} e^{-\rho a} [v(l(t,a,\theta)) + \eta(a,\theta)] da$$

The parameter λ captures the intergenerational discount rate in the social welfare function. The planning problem is then to maximize the value of the objective in (34) over the set of feasible and incentive-compatible allocations. Note that since $r = \rho$, any solution to this problem must prescribe constant consumption over the life-cycle for each individual.

The planning problem in the main text, alternatively, maximizes the value for one generation subject to incentive compatibility and a feasibility constraint of the form

$$H + \int_0^{\bar{T}} \int_{\Theta} e^{-rt} c(t,\theta) dF(\theta) dt = \int_0^{\bar{T}} \int_{\Theta} e^{-rt} \varphi(t,\theta) l(t,\theta) \mathbf{1}[t \leq R(\theta)] dF(\theta) dt$$

Label the solution to this problem $\{c^*(t,\theta;H), l^*(t,\theta;H), R^*(\theta;H)\}$. We refer to the planning problem associated with a single generation as **P1** and the planning problem in the OLG economy as **P2**. Let the value of the social welfare in **P1** be given by $U^*(H)$.

For the steady state of **P2**, let the allocations be given by $\{c_{ss}(a,\theta), l_{ss}(a,\theta), R_{ss}(\theta)\}$. In addition, define H_{ss} as the difference between the present value of labor earnings and the present value of consumption in the steady state. Our main claim is that the steady state allocation coincides with the solution to **P1** for H_{ss}. The idea is that otherwise we can simply replace the steady state allocation with the solution to **P1** for generations that are born late in time.

Claim 1 *Given a unique solution to **P1**, the steady-state allocation of **P2** coincides with that of the one generation economy with H_{ss}, i.e.,*

$$c_{ss}(t,\theta) = c^*(t,\theta;H_{ss}), l_{ss}(t,\theta) = l^*(t,\theta;H_{ss}), R_{ss}(t,\theta) = R^*(t,\theta;H_{ss}).$$

Proof. Suppose not. Consider the solution to **P2**, and let $H(t)$ be the difference between the present value of labor earnings and consumption for generation t. Given the definition of the steady state of **P2**, it must be that $H(t)$ converges to H_{ss} as t tends to ∞. Since the solution to **P1** is unique and the steady-state allocation does not coincide with it, it must be that social welfare of the steady-state allocation is less than $U^*(H_{ss})$. Since $H(t)$ converges to H_{ss} and $U^*(H)$ is continuous in H, for large enough t , the aggregate welfare of generations born at t is less than $U^*(H_t)$. Replace these allocations with the solution of **P1** for $H = H(t)$. These allocations are feasible since the difference in present value of consumption and labor earnings remains constant. Furthermore, they are incentive compatible and they deliver a higher value of social welfare. This implies that the initial allocation cannot be a solution to **P2**. ■

C Further details for Sections 4 and 5

We expand here on the main text to provide further details of the sample construction, the estimation, and the sensitivity checks of the results.

Data sets. We use two sources of longitudinal individual-level data, the HRS and the PSID. To work with the HRS, data we use version K of RAND HRS files, which are based on all publicly available surveys from 1992 to 2008. The details of cleaning, processing, and consolidating raw HRS variables are extensively documented in RAND HRS version K documentation available online from RAND Corporation.

Raw data files of the PSID contain a well-known range of inconsistencies, impossible answers, and other issues. We use the data set from Heathcote, Perri, and Violante (2010), who carefully address these issues, and refer to that paper for details. We use Sample A of Heathcote, Perri, and Violante (2010), which is their most inclusive sample and is essentially a cleaned version of the data that come from the Survey Research Center (SRC) sample of the PSID using all of the annual surveys from 1967 to 1996 and the biennial surveys for 1999, 2001, and 2003. One issue particularly well-known in the literature, potentially important for our purposes, is top coding of incomes in the PSID. Heathcote, Perri, and Violante (2010) address it by fitting a Pareto distribution in the right tail. For our baseline we take this data with the fitted Pareto tail. Among the robustness exercises below, we check the effects of removing the fitted thick right tail to make sure it does not introduce artificial qualitative features into our results.

Sample selection. As a baseline we use males of one U.S. cohort referred to as 1940 cohort. It includes males born between 1931 and 1941. This coincides with the initial HRS

Table 3: Distribution of annual hours worked, by age.

Age	Annual hours					
	[0,500)	[500,1000)	[1000,1500)	[1500,2000)	$\geq$2000	hours/worker
60	0.03	0.04	0.06	0.17	0.70	2150
61	0.04	0.03	0.06	0.17	0.66	2124
62	0.08	0.05	0.07	0.18	0.57	2015
63	0.10	0.09	0.08	0.13	0.52	1909
64	0.12	0.09	0.12	0.12	0.45	1789
65	0.11	0.10	0.14	0.14	0.43	1777
66	0.16	0.09	0.16	0.13	0.33	1642
67	0.25	0.08	0.12	0.08	0.25	1605
68	0.26	0.08	0.14	0.09	0.20	1539
69	0.28	0.08	0.11	0.09	0.19	1538
70	0.24	0.12	0.15	0.09	0.19	1418

Note: 1940-cohort males in the pooled sample of the HRS and the PSID.

cohort, which was first interviewed in 1992 and subsequently every two years, providing the longest observed cohort over time in the HRS. This also implies that the cohort have approached age 60 by the year 2000 and hence we use a stylized version of the U.S. Social Security system from 2000.

This leads us to start with 9,638 individuals in the PSID and 15,959 in the HRS. After we restrict the age to at least 20 and gender to males, we obtain in the PSID sample 6,918 and in the HRS sample 3,656 individuals. When we restrict observations per individual to be at least 23 in the PSID data (later we check sensitivity by relaxing this restriction), we are left with a sample of 1,116 individuals. We restrict observations per individual to at least 5 in the HRS to get 971 individuals. The final count of individuals in the pooled HRS-PSID data set is 2,087.

In terms of observations, we start with 93,924 in the PSID and 56,667 in the HRS. After restricting the age and gender, we obtain 77,157 observations in the PSID and 26,602 observations in the HRS. Restricting observations to at least 23 per individual in the PSID leaves 30,751 observations; restricting observations to at least 5 per individual in the HRS leaves 5,788 observations. The final count of observations in the pooled HRS-PSID data set is 36,539, providing on average 18 observations per individual.

Life span, retirement ages, and Social Security claiming ages. Each individual enters the quantitative environment at age 20. We need three additional ages for each individual: the age when the individual claims Social Security benefits, $S(\theta)$; when the individual retires, $R(\theta)$; and when the individual dies, $T(\theta)$.

As a baseline, we take $\bar{T}$ to be 81.6 from the Social Security Administration's Life

Table 4: Distributions of retirement ages, benefit claiming ages, and life spans, by earnings decile.

	Earnings percentile		
	Bottom decile	Median	Top decile
Retirement age by definition:			
baseline	67.6	66.9	68.5
alternative	66.3	66.1	66.0
Retirement age by sector:			
manufacturing and mining	67.2	66.3	64.8
retail	67.8	70.3	68.0
professional services	70.1	66.8	70.5
Retirement age by education:			
less than high school	67.7	67.0	67.5
high school and above	67.1	66.8	68.7
Social Security claiming age	62.6	63.4	63.5
Life expectancy at age 60	78.3	81.4	83.8

Note: 1940-cohort males in the pooled sample of the HRS and the PSID. Percentiles are constructed as in the main text.

Tables in Bell and Miller (2005) for 1940-born males (see also our discussion above of the comparison with the steady-state of an overlapping generations economy). Since the retirement behavior is part of the focus here, we take the life span for the individuals who survived at least until age 60.

We obtain retirement ages $R(\theta)$ by using two definitions, a baseline and an alternative. Our baseline definition of retirement uses the RAND HRS variable *RwLBRF*, which aims at consolidating all available in the HRS sources of information about the individual's labor-force status. The consolidation uses a variety of questions, depending on the wave, listed in the online RAND files documentation. The evidence of working hence comes from multiple sources reconciled and summarized in *RwLBRF*. Of importance here is that this consolidation aims at separating retirement from claiming pension benefits from Social Security, unemployment, partial retirement, or claiming to be retired while also reporting labor earnings.

We first identify the first and last non-missing labor-force status values for an individual according to *RwLBRF*. Then, within this range, only if we observe a change in status from a non-missing non-5 value to the value of 5 in two consecutive waves (to retired from any other status), we define the person as potentially retired in the latter wave. Finally, the maximum among potential retirement ages is used as the baseline retirement age.

We check the implications of alternative definitions of retirement in the data. Retire-

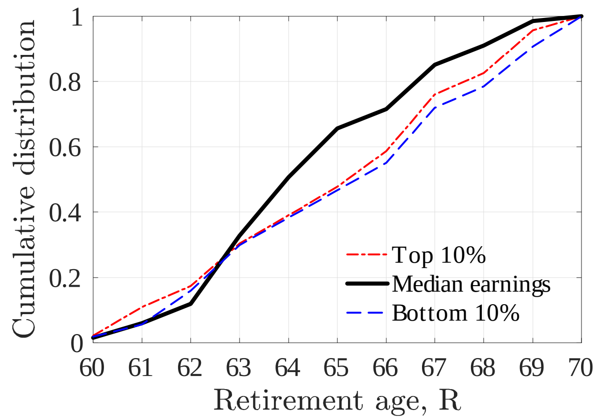

Figure 13: Cumulative retirement age distribution for selecred percentiles of lifetime earnings.

ment in our setting means no work hours (excluding unemployment) from a given age onward. As a check, we use a definition based purely on hours worked reported in the PSID and the HRS, referred to as the alternative definition in the main text. Following Guvenen (2009), we define individuals as retired if hours worked fall below 520 permanently (i.e. 10% of 5,200 hours, the likely highest sustainable annual hours according to Guvenen (2009)). The Pearson correlation coefficient between retirement age according to the baseline definition and this alternative definition is 0.67. The average retirement ages by the two definitions in the pooled sample are much less than a standard deviation apart.

The individual information about Social Security benefits claiming ages $S(\theta)$ (as opposed to retirement ages) is summarized in the RAND HRS in the variable *RASSAGEB*. We adopt it as the baseline definition of Social Security claiming age in the quantitative analysis.

Distribution of hours and retirement ages. Here we take a more detailed look at the behavior of hours worked around the time of retirement and then describe how the retirement ages vary with the definition of retirement, with education, by sector, and with labor earnings.

First, we examine the labor supply behavior around retirement to see if in our sample retirement appears to be associated with smooth transitions or with more abrupt changes from full-time work to not working. For instance, Rogerson and Wallenius (2013) show evidence of quite abrupt transitions for the CPS and the PSID separately as well as review the evidence in the literature based on the HRS. Table 3 suggests similar behavior in our pooled sample of the HRS and the PSID. At age 60, by far the most common average amount of hours worked per year is greater than 2,000. Already starting at age 67 and

Table 5: Statistical models of productivity-age profiles.

Parameter	Model 1	Model 2	Model 3
a	0.126171 (0.0309)*	0.213422 (0.0732)*	0.128442
β_2	-0.00172 (0.000233)*	-0.00214 (0.000263)*	-0.00172
β_3	7.267x10^{-6} (1.723x10^{-6})*	0.00001 (1.911x10^{-6})*	6.5403x10^{-6}
Individuals:	2,087	2,087	N/A
Observations:	36,539	22,822	N/A

Note: * Indicates significance at the 1 percent level.

quite clearly by age 70, the most common hours worked are less than 500, with an increase in that group from 16% to 25% at age 67.

Table 4 summarizes how the distribution of retirement ages changes with the definition of retirement, with education, and by sector. The baseline and the alternative definitions of retirement result in essentially similar retirement age patterns. More physically demanding sectors, such as manufacturing and mining, deviate from the general pattern, with retirement ages that decline with higher earnings. Individuals in less physically demanding sectors tend to retire at older ages, both at the bottom and at the top of the earnings distribution. More educated individuals tend to retire at younger ages throughout the distribution of earnings. Using the baseline definition of retirement, Figure 13 compares the cumulative distribution of retirement ages (retirement hazard ratios) of the median earnings decile with that of the top and the bottom deciles, summarizing the overall distributional differences.

Productivity-age profiles. The statistical model described in the main text can be more explicitly written as

$$w_{it} = \left[\beta_0 + \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 x_{it}^3\right] + \sum_{j=2}^{I} \gamma_{0i} \left(d_j\right) + \sum_{j=2}^{I} \gamma_{1i} \left(d_j x_{it}\right) + \varepsilon_{it},$$

where w_{it} is the logarithm of effective reported labor earnings per hour for person $i = 1,..,I$ at time t, x_{it} is a measure of experience, taken to be age in the baseline, so that the term in brackets is the logarithm of the common age component, d_i are individual dummies, and the constraints implied by our parametric assumption is given by $\beta_1 = a\beta_0$ and $\gamma_{1i} = a\gamma_{0i}$ for $i = 2,..,I$. That is, we estimate non-linear equations

$$w_{it} = \left[(1 + a x_{it})\,\beta_0 + \beta_2 x_{it}^2 + \beta_3 x_{it}^3\right] + (1 + a x_{it}) \sum_{i=2}^{I} \gamma_{0i} \left(d_i\right) + \varepsilon_{it},$$

Table 6: Summary statistics for earnings deciles used in constructing productivity-age profiles.

Earnings Decile	Average Decile Statistic			
	Annual Hours	Education (years)	Married (fraction)	Caucasian (fraction)
1st (bottom)	2227.19	11.0819672	0.85	0.85
2nd	2274.84	11.7015504	0.89	0.82
3rd	2217.93	12.0788382	0.86	0.89
4th	2304.49	12.4729730	0.93	0.92
5th	2237.65	12.8357488	0.92	0.91
6th	2175.42	13.2620321	0.88	0.93
7th	2123.48	13.6042781	0.91	0.93
8th	2160.45	14.0960452	0.88	0.94
9th	2066.72	14.5444444	0.82	0.94
10th (top)	2077.16	15.5217391	0.86	0.97

Note: deciles are based on the predictions from the baseline estimation.

where $\{a, \beta_0, \beta_2, \beta_3, \gamma_{0i}\}$ are estimated using GMM, reported in the Model 1 column in Table 5.

Given estimates $\left\{\hat{a}, \hat{\beta}_0, \hat{\beta}_2, \hat{\beta}_3, \{\hat{\gamma}_{0i}\}_{i=2}^{I}\right\}$, we construct a vector of estimates $\{\log\hat{\theta}_i\}_{i=1}^{I}$ defined as

$$\{\hat{\beta}_0, \hat{\beta}_0 + \hat{\gamma}_{02}, \hat{\beta}_0 + \hat{\gamma}_{03}, \cdots\}$$

so that the predicted log productivity-age profiles are given by

$$\log\hat{\varphi}_t^n = (1 + \hat{a}x_t)\log\bar{\theta}_i^n + \hat{\beta}_2 x_t^2 + \hat{\beta}_3 x_t^3$$

where $\log\bar{\theta}_i^n$ is the average of $\log\hat{\theta}_i$'s for a group $n = 1..N$ and $x_t = 20..80$ is age. That is, the individual fixed effects are interpreted as individual type and $\log\varphi$ is proxied with the logarithm of effective labor earnings per hour, i.e., the computed ratio of all labor earnings to total hours reported in the PSID and the HRS. The labor earnings are taken directly from our pooled sample variables, converting to constant 2000 dollars. We aim to capture total labor earnings summing over the variables containing "salaries and wages", "separate bonuses", "the labor portion of business income", "overtime pay", "tips", "commissions", "professional practice or trade payments", and "miscellaneous labor income". We similarly capture the total reported hours.

For our baseline, we set $N = 10$ where each group is defined to be a decile, later changing the number of groups to $N = 5$ and $N = 20$. Overall, the Pearson correlation coefficient between estimated type $\hat{\theta}$ and the average annual earnings is 0.85. The same correlation with average annual earnings after age 60 is 0.80. To a first approximation,

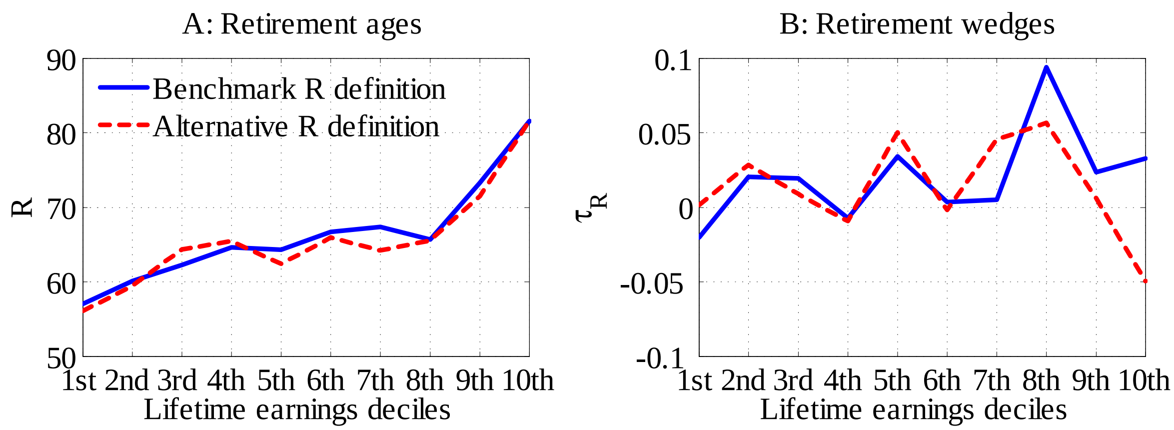

Figure 14: The effects of the variation in the definition of retirement on efficient retirement ages as a function of lifetime earnings (Panel A) and their associated optimal retirement wedges (Panel B).

the deciles are also referred to as average annual-earnings deciles. Table 6 reports the analogues of the summary statistics in the main text for each of the baseline deciles.

Estimation sensitivity checks. We implement several variants of the above estimation approach resulting in qualitatively virtually indistinguishable simulations from those reported in the main text. First, we varied the number of groups to $N = 5$ and $N = 20$ with virtually identical results. Then, instead of age, we also used two alternative measures of experience x_{it}:

$$x_{it}^1 = age_{it} - 19$$

and

$$x_{it}^2 = age_{it} - \max\{edu_{it}, 10\} - 5.$$

We find that replacing age with a measure of experience accounts for the rightward shift in the peak of the profile as the type increases, as is well-known in labor literature. That is, using x_{it}^1 instead of age produces very similar productivity profiles to the ones reported in the main text, except the peaks of the profiles of different types are much closer to each other, with x_{it}^2 resulting in virtual alignment. The latter variant is also very close to the profiles we obtained by closely following the approach of Nishiyama and Smetters (2007) and grouping individual observations by type into one of seven bins, each for a 10-year interval of ages – 25-35 years old, 34-45 years old, ..., 74-85 years old (the few remaining individuals older than 85 were put in the last group) – and extrapolating by using shape-preserving cubic splines to obtain the complete productivity-age profiles.

As another check of the sensitivity of the estimation results, we relaxed the require-

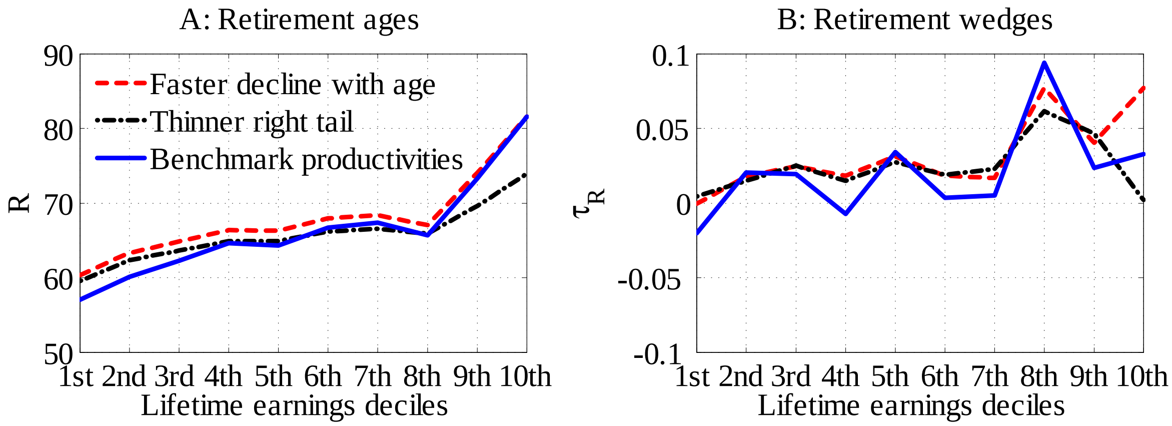

Figure 15: The effects of the variation in the estimates of the productivity-age profiles on efficient retirement ages as a function of lifetime earnings (Panel A) and their associated optimal retirement wedges (Panel B).

ment that in the pooled sample the individuals from PSID have 23 observations, and replaced it with 5 observation requirement, as used for the HRS individuals. To keep the GMM estimation computationally feasible, we randomly selected the same number of individuals as in the original sample so that the total pooled-sample size remained the same. The resulting estimates are reported in Table 5 in the Model 2 column. The estimates of the baseline Model 1 are within two standard deviations from Model 2.

Finally, given a debate in the literature about how much of the curvature in the profiles is driven by the parametric choices, especially at later ages, we study the effects on the constrained optimum of the following check: we force the curvature to vary by exogenously increasing parameter a by 20 percent and reducing parameter β_3 by 10 percent. The resulting parameters are reported in Table 5 in the Model 3 column and the constrained optimum is simulated below with very similar results to the baseline.

Further robustness results. In addition to the analysis of the sensitivity of the normative simulation results in the main text, we provide here three further sets of deviations from the baseline simulation. First, we show the robustness of the main normative simulation results with respect to the definition of retirement ages in the data. As there is no single universally accepted definition in the literature, we adopted two definitions we judged to be potentially as different as the data would allow. The definitions are discussed in the main text, with additional details above. Figure 14 illustrates the resulting comparison to the baseline simulated constrained optimum. As the main text explains, it is reasonable to focus on the profiles of retirement ages and on optimal retirement wedges since the rest of the constrained optimum follows. Figure 14 reveals qualitatively only mi-

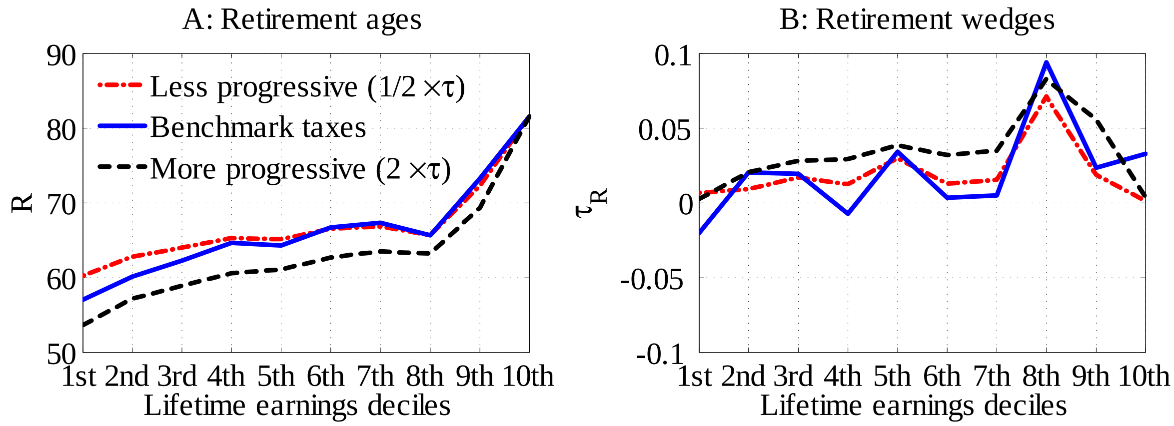

Figure 16: The effects of the variation in the estimates of the effective status-quo policies on efficient retirement ages as a function of lifetime earnings (Panel A) and their associated optimal retirement wedges (Panel B).

nor differences that would lead one to make the same conclusions as in the main text.

Second, a potentially key deviation from the baseline simulation results may be introduced by the parametric restrictions on the productivity-age profiles as we discussed above. To explore these effects, we varied our assumptions in the estimation as discussed in the previous section. Figure 15 compares the baseline simulation of the optimum to two of the most extreme cases we found. Once again the differences appear to be quantitatively minimal.

Finally, we explore how robust our main quantitative insights are with respect to the estimates of the effective tax functions. Even though the estimates in the main text follow the literature, they are at the core of the estimated fixed costs and thus it is instructive to illustrate how sensitive the simulated constrained optimum is to possible errors in the estimates. Figure 16 compares the results from the baseline calibration to a calibration based on a tax function with the progressivity parameter arbitrarily forced to be twice as large and half as large as the estimate from the literature that we used in the baseline and discussed in the main text. As with all of the cases above, this results in quantitatively minimal changes and suggests the same general conclusions we discussed in the main text.